DEVELOPMENT, VALIDATION, AND APPLICATIONS OF THE PEPSAVI-MS PIPELINE FOR NATURAL PRODUCT BIOACTIVE PEPTIDE DISCOVERY

Christine L. Kirkpatrick

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry in the School of Arts and Sciences.

Chapel Hill
2018

Approved by:

Leslie M. Hicks

Gary L. Glish

James W. Jorgenson

Bo Li

Elizabeth A. Shank

ABSTRACT

Christine L. Kirkpatrick: Development, validation, and applications of the PepSAVI-MS pipeline for natural product bioactive peptide discovery
(Under the direction of Leslie M. Hicks)

The recent increase in multidrug-resistant pathogens and associated morbidity/mortality demonstrate the immediate need for new antibiotic backbones with novel mechanisms of action. While natural products are a well-studied source of biologically active small molecules, peptidyl factors contributing to their medicinal properties remain largely unexplored. To expedite the search for this exciting class of compounds this dissertation describes the development, validation, and applications of PepSAVI-MS (Statistically-guided bioactive peptides prioritized via mass spectrometry). This highly versatile pipeline employs whole-cell bioactivity screening coupled with mass spectrometry and bioinformatics to identify bioactive peptides from complex biological extracts (Chapter 2). The development and validation of PepSAVI-MS was originally described using a botanical source through the successful detection and identification of a known antimicrobial peptide, cycloviolacin O2 (cyO2), from *Viola odorata* (Chapter 3). In addition to pipeline validation, this study widened the known antimicrobial spectrum for *V. odorata* cyclotides, establishing antibacterial activity of cyO2 against *Acinetobacter baumannii*, and explored novel anticancer activities for cycloviolacins by their cytotoxicity against ovarian, breast and prostate cancer cell lines.

To demonstrate the versatility and wide applicability of PepSAVI-MS, validation studies were performed using fungal (Chapter 4) and bacterial sources (Chapter 5), with optional

experimental modifications to highlight pipeline utility. Identification of the virally-encoded antifungal killer toxin KP4 from *Ustilago maydis* P4 and the bacteriocin Bac-21 from *Enterococcus faecalis* pPD1 demonstrates proof-of-principle for bioactive peptide discovery from bacterial and fungal secretomes.

Using the validated pipeline, a variety of natural product sources are probed to prioritize highly active species for downstream analysis. The plant species *Amaranthus tricolor* is presented as an example of novel antimicrobial activity identified with PepSAVI-MS (Chapter 6). As demonstrated herein, the developed pipeline is powerful and highly versatile; PepSAVI-MS can incorporate material from a variety of natural product sources and is able to accommodate any developed bioactivity screen. The flexible nature of this pipeline presents a valuable tool for diverse natural product exploration, and has the potential to lead to the discovery of a wealth of bioactive peptides.

To my family.

ACKNOWLEDGEMENTS

Reflecting on my time at UNC, there are so many people that have had an impact on my journey, and for which I will be forever grateful. First, I want to thank my family for all of their love and support. I could not have continued on this journey without your encouragement. To my parents, thanks for always being there for me and being my biggest fans. To my sisters, thanks for putting up with me and giving me a place to escape when grad school became too much. To my friends and extended family, thank you for supporting me even when you knew it meant that I would be missing countless holidays, birthdays, and life adventures. I promise I will make it up to you! And to Derek, thank you for being my rock throughout this process. I could not have done this without your constant love and support, and can't imagine going through this with anyone else.

To the Hicks Lab, I can't thank everyone enough for walking this journey with me. Leslie, thank you for your constant support and guidance. You have constantly pushed me to reach my full potential, and have inspired me in countless ways. I have learned so much about being a great scientist, mentor, and leader. I will be forever grateful for the opportunity to be one of your first graduate students. To my lab mates, we have been through so much together. We have celebrated countless publications, grants, and milestones, and commiserated through failed experiments and grad school troubles. I would not have wanted to share these experiences with anyone else. Emily, thank you for being my partner in crime and my best friend throughout it all. I guess I'll come visit you after all of this. Evan, thanks for your patience with all my stats

questions and for the many memories through the years. To Nicole and Tessa, keep killin' it with bioactive peptides. I am excited to see all the things you will do! To Megan and Tony, hold down the fort when we are gone. And to all of my undergraduates along the way, thanks for your hard work and comic relief. Keep following your dreams – I am proud of you all!

And to my friends, thank you, thank you for the wonderful memories. We have been through it all - from late nights studying, to hours spent in electronics lab, to beach vacations, to a basketball national championship. To Emily, Rima, Alex, Micah, Nate, Jake, Boyce, and Matt, I am so glad we went through this journey together. Thanks for always being there to celebrate achievements, commiserate failures, and to provide endless laughter and friendship. I will miss you all, but I know we will all go on to do great things and can't wait to see you in the future. It's a great day to be a Tar Heel.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACPP	Activity-correlated quantitative proteomics platform
AMPs	Antimicrobial peptides
AMR	Antimicrobial resistance
ATCC	American type culture collection
BCP	Bacterial cytological profiling
CETSA	Cellular thermal shift assays
cyO2	cycloviolacin O2
Da	Daltons
ED	Effective dose
Erm	Erythromycin
ESKAPE	*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* species
GAFFI	Global Action Fund for Fungal Infections
GUI	Graphical user interface
HPLC	High performance liquid chromatography
IC	Inhibitory concentration
IDA	Information dependent acquisition

LC-MS	Liquid chromatography-mass spectrometry
m/z	Mass to charge ratio
Mcl-1	Myeloid cell leukemia 1
MDR	Multi-drug resistance
MHB	Mueller Hinton broth
MIC	Minimum inhibitory concentration
MM	Minimal media
MOAs	Mechanisms of action
MPA	Mobile phase A
MPB	Mobile phase B
MS	Mass spectrometry
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
MW	Molecular weight
N2(l)	Liquid nitrogen
OD	Optical density
PepSAVI-MS	Statistically guided bioactive peptides prioritized via mass spectrometry
PRM	Parallel reaction monitoring
PTMs	Post-translational modifications

PVPP	Polyvinylpolypyrrolidone
RFU	Relative fluorescence units
RiPPs	Ribosomally-synthesized, post-translationally modified peptide natural products
RP-HPLC	Reversed phase- high performance liquid chromatography
SCX	Strong cation exchange
SSPS	Solid-phase peptide synthesis
TIC	Total ion chromatogram
UM	*Ustilago maydis*
UMP6	*Ustilago maydis* carrying the P6 virus
WCX	Weak cation exchange

CHAPTER 1: Introduction

1.1 The Threat of Antimicrobial Resistance

Antimicrobial resistance (AMR) is present in every country worldwide and has a direct impact on global health and economy. Human fatalities associated with antimicrobial resistance are estimated at 10 million per year worldwide and health-related economic losses are estimated to approach an ~3% decrease in gross domestic product exceeding $100 trillion by 2050[1]. Treating common bacterial and fungal infections is becoming increasingly challenging as new resistance mechanisms develop and current methods are accelerated by the misuse of antimicrobial drugs[1].

Despite notable success of antibiotics in the past 70 years, bacterial diseases remain a leading cause of mortality worldwide, killing over 17 million people each year[2-8]. In the US, there are two million hospital acquired infections each year, resulting in ~100,000 deaths[9]. Almost a quarter of this number is caused by organisms for which there are no treatment options remaining[10]. Specifically, the emergence of antibiotic resistance in ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* species) pathogens has led to a significant increase in difficult-to-treat nosocomial infections with associated increases in morbidity/mortality. With many bacteria now unresponsive to conventional therapeutics and the anticipated trend toward increasing broad-spectrum resistance by 2030, there is an undeniable and desperate need to develop new antimicrobial therapeutics.

Antimicrobial illness and resulting deaths are not only caused by bacteria, but fungi as well. Fungal infections are the fourth most common illness on Earth (GAFFI report, 2018). While

fungal infections are largely treatable in healthy patients, infections in immunocompromised patients are often deadly. The Global Action Fund for Fungal Infections (GAFFI) reports that 150 people die every hour from fungal infections worldwide, and that these infections reach over 80% of the global population. In addition to this burden on the global population, fungal pathogens are one of the leading causes of agricultural crop loss, accounting for >12% of annual crop loss worldwide[11]. Among the most common and relevant agricultural fungal pathogens are *Fusarium* sp., responsible for Fusarium Head Blight in wheat and barley[12], and *Ustilago maydis*, responsible for corn smut virus[13]. With the threat of both bacterial and fungal pathogens, it becomes imperative that new methods and innovative technologies are employed to replenish our dwindling arsenal of antimicrobial drugs.

1.2 The Scarcity of the Drug Discovery Pipeline

While >100 antibacterial agents have been developed to treat human disease since the discovery of sulfonamides, our current therapeutic options are worse now than at almost any time since the discovery of antibiotics[14]. Current methods for antibiotic drug discovery are being outpaced by the rise of antimicrobial resistance[8], and the recent detection of plasmid-borne *mcr*-1 mediated colistin resistance in *E. coli* within the United States further emphasizes the dire need for antimicrobial discovery with novel mechanisms of action[15].

Successful drug development reaching commercialization over the last fifty years has largely focused on the discovery of natural antibiotic compounds followed by incremental changes to chemical structure of these backbones[16]. Advances in large-scale genome sequencing was predicted to herald an era of genes-to-drugs, where specific bacterial targets would be identified and a small molecule inhibitor would be found through screening of large chemical libraries to target that specific gene product[17]. Contrary to expectations, genome sequencing has revealed

bacterial physiology to be more complex than anticipated, with many genes having unknown functions. Characterized pathways are extraordinarily complex, contain redundancy, and/or have the capacity to rapidly develop resistance to new compounds. Also, unanticipated challenges have arisen in the search for appropriate targets for small molecule inhibitors[3]. This is further complicated by difficulties associated with bacterial cell wall and membrane permeability relative to host bioavailability[3]. Attempts to target a specific, isolated gene product fail to identify compounds that have the physiological capacity to inhibit growth of, or kill, intact cells. Attempts at using whole cell screening assays are not high throughput as they test individual compounds against a bacterial, fungal, or cancer cell culture. As a consequence, large pharmaceutical companies have retreated from antibiotic research and smaller biotech companies that are still pursuing this research are returning to low efficiency natural product isolation. A highly efficient screening pipeline that tests antibiotic-like compounds in complex mixtures against growing bacterial cells is needed to overcome the difficulties seen with the current discovery pipelines[3, 16-17].

1.3 Peptides as Drugs

The push for alternative strategies to increase identification of novel lead compounds has contributed to increased interest in the development of peptides as drug candidates. Peptide-based therapeutics boast higher selectivity than small molecules and offer better tissue penetration and lower manufacturing costs than proteins[18]. While peptides were considered as potential therapeutics without success many years ago, increasing prevalence of antibiotic resistance and the steady entry into the "post-antibiotic era" encourage reconsideration of this class of compounds. In fact, several peptide antibiotics (i.e. polymyxin B, colistin, daptomycin) are listed in the WHO Model List of Essential Medicines as Group 3 – Reserve Group

Antibiotics that should be treated as options of "last resort" when all alternatives have failed in order to preserve their effectiveness[19]. Recent advances in peptide modification, formulation, and delivery methods can address known limitations[18], including modification of peptide length/content to increase selectivity[20-23], stapling and/or peptidomimetic conversion techniques to improve pharmacokinetic properties[24-26], incorporating active peptides into stable scaffolds capable of oral delivery methods[27], and encapsulation to protect from metabolic degradation or allow for tissue-specific release[23, 28]. A recent study reported that stapling of the myeloid cell leukemia 1 (Mcl-1) protein not only increased stability of the α-helical structure but also increased target binding affinity and specificity for Mcl-1 over related Bcl-2 proteins, further highlighting the potential of this class of compounds[29]. Bioactive peptides also hold great potential in combination therapy-based applications as their cell penetration mechanism both initiate cell death and allow for increased penetration of drugs with intracellular targets. The advantages offered by peptide-based drugs have led to the recent approval and additional development of a number of compounds, with approximately 140 peptide-based compounds currently in clinical trials and an estimated 500-600 additional drugs in development[18,30].

1.4 Natural Products as a Source of Novel Lead Compounds

Natural products play a pivotal role in drug discovery yet remain a largely untapped resource. Plant- and microbial-based natural products have long been sources for lead compounds and inspiration for new chemistries at the forefront of drug discovery[31-38].

Plant-based natural products have been relied upon for centuries for primary health care in developing nations and have contributed substantially to Western medicine. While tens of thousands of plant species are estimated to be used traditionally, only a small percentage have been screened for biological activity[37-40]. Lacking the somatic adaptive immune system found in

mammals, plants rely on a complex innate immune response for protection against invading fungal and bacterial pathogens[41], including secretion of peptide "danger signals" with broad-spectrum antimicrobial properties[42-46]. Despite hundreds of millions of years of antimicrobial peptide (AMP)-modulated host-pathogen interactions, antimicrobial resistance to continually evolving plant AMPs has seldom been seen in nature, demonstrating the adaptability and efficacy of these defense molecules[47].

The primary and tertiary structures of known AMPs are highly diverse- however, most are <10 kDa and require a net positive charge and intact disulfide bridges for bioactivity. Significant variations in tertiary structure have led to the classification of seven different plant AMP families (cyclotides, thionin, snakin, defensin, knottin, hevein-like and lipid-transfer; Figure 1.1), all of which are involved in innate pathogen defense[48-49]. Of the six kingdoms represented in the antimicrobial peptide database, plants represent the second most prevalent source of AMPs (behind animals)[50].

From the discovery of the first plant AMP over 45 years ago, a purothionin in wheat flour (*Triticum aestivum*) with gram-positive antibacterial activity, AMPs have been isolated from virtually every plant tissue (seeds, roots, stems, leaves, and flowers) and from a variety of botanical species as diverse as grain crops, hellebores, tobacco, euphorbias, violets, coffee, and *Arabidopsis thaliana*[45, 49, 51-52]. Snakins (i.e. Snakin-1 from *Solanum tuberosum* and Snakin-2 from *Solanum lycopersicum*) are six intramolecular disulfide-containing AMPs that exhibit broad-spectrum activity against bacteria and fungi in low micromolar concentrations[53-54]. Plant defensins are ubiquitous in the plant kingdom and analogous to defensins found in animals, fungi, and insects. They display potent insecticidal, antifungal, and antibacterial activities likely derived from diverse mechanisms of action (MOAs), including inhibition of protein translation,

α-amylases and proteases in target cells, electrostatic and amphipathic membrane interactions, and ion channel blocking[55]. Cyclotides and knottin-like peptides (i.e. cycloviolacins 1-25 from Viola sp.) are structurally robust molecules, containing three interlocking disulfides, with high levels of antifungal, insecticidal and antibacterial activities[56]. Despite this impressive display, the plant kingdom as a whole has barely been tapped with respect to depth of AMP discovery within a given species or across the >300,000 known plant species. As such, plants remain a rich reserve of uncharted bioactive peptides.

While lacking the complexity of the plant immune system, microbes produce a wide variety of defense molecules to leverage competitive advantage over invading microbial pathogens. Both bacteria and fungi have developed intricate systems to for environmental protection, including the production of bacteriocins in bacteria[57] and the establishment of symbiotic relationships with viruses in fungi[13]. Although created for survival, the innate bioactivity of these peptides poses a promising source of molecules to exploit for drug discovery, agriculture, and food safety applications[58-59]. The antimicrobial peptide, copsin, from the basidiomycete *Coprinopsis cinerea*, was first discovered in 2014 and has potent activity against gram-positive bacteria including both the human pathogen *Enterococcus faecium* and the foodborne pathogen *Listeria monocytogenes*[60]. Additionally, there has been a growing interest in the agricultural use of bioactive peptides for crop protection, including transgenic expression in plants and topical application as biopesticides. Spear-T, a bioinsecticide derived from spider venom and marketed by Vestaron, has recently been approved by the EPA for commercial use. Importantly, unlike many currently used neonicotinoid-containing products, Spear-T has no adverse effects on bees or other beneficial insects, highlighting additional benefits of certain peptide-based agricultural products[61]. Fungi-microbe interactions have also inspired the use as these natural peptides in

food safety applications, including the use of bacteriocins from lactic acid bacteria as starter cultures in food fermentation[62].

1.5 Current Methods for the Discovery of Bioactive Peptides

Traditional proteomics/peptidomics profiling aims to identify and characterize all peptide and protein constituents in an organism. While essential for in-depth understanding of an organism, targeted methods to identify peptidyl constituents are necessary to facilitate rapid bioactive peptide discovery. Common approaches for targeted bioactive peptide identification include bioassay-guided fractionation and genome mining.

1.5.1 Bioassay Guided Fractionation

Bioassay-guided fractionation approaches employ iterative rounds of fractionation, bioactivity screening, and LC-MS analysis to isolate a single bioactive compound (Figure 1.2). Fractions typically undergo 3+ rounds of fractionation to isolate the activity to a single compound, with each iterative round of fractionation resulting in loss of material and inefficiency of process. As such, this approach is labor intensive and displays an innate bias towards highly active and abundant compounds, often resulting in isolation and re-identification of already known compounds[63-67]. However, bioassay guided fractionation has served as the gold standard for bioactive compound discovery as it is simplistic in nature, provides a direct measure of bioactivity, and is amenable to any developed bioassay.

1.5.2 Genome Mining

Alternatively, genome mining can be used to predict novel bioactive peptides in a given sequenced organism (Figure 1.3)[68]. Genomic prediction relies on a strong knowledgebase of known bioactive peptide sequences, processing enzymes, and biosynthetic pathways to accurately predict potential bioactive peptides. A variety of computational algorithms (e.g.

Pep2Path[69], RiPPquest[70], and NRPquest[71]) have been developed to assist in identifying target biosynthetic gene clusters. Once identified, biosynthetic gene clusters capable of producing bioactive peptides can be cloned and expressed in a host organism to allow for access to bioactive peptides that are not constitutively expressed in a given organism. Although this approach makes previously inaccessible peptide products now accessible, it requires a strong training data set of known bioactive peptide biosynthetic pathways. While highly useful in bacterial species, these systems are less developed in complex organisms including plants. Furthermore, this approach has no direct measure of bioactivity and requires extensive follow-up experiments for activity validation. Low throughput, replication of known compounds, and inefficiency of process result in either case[65, 72-73].

1.6 The PepSAVI-MS Pipeline

To expedite the search for natural product bioactive peptides, the PepSAVI-MS pipeline has been developed (Chapter 2). Relying on a multi-pronged approach, the PepSAVI-MS pipeline utilizes selective extraction and fractionation of peptide source material, bioactivity screening, and statistics-guided mass spectrometry-based peptidomics for the targeted identification and characterization of only putative bioactive compounds. The initial development and validation of PepSAVI-MS is described using a botanical source through the successful detection and identification of a known antimicrobial peptide, cycloviolacin O2 (cyO2), from *Viola odorata* (Chapter 3). To demonstrate the versatility and wide applicability of PepSAVI-MS, validation studies were performed using fungal (Chapter 4) and bacterial sources (Chapter 5), with optional experimental modifications to highlight pipeline utility. Identification of the virally-encoded antifungal killer toxin KP4 from *Ustilago maydis* P4 and the bacteriocin Bac-21 from *Enterococcus faecalis* pPD1 demonstrates proof-of-principle for bioactive peptide discovery

from bacterial and fungal secretomes. Using the validated pipeline, a variety of natural product sources are probed to prioritize highly active species for downstream analysis. The plant species *Amaranthus tricolor* is presented as an example of novel antimicrobial activity identified with PepSAVI-MS (Chapter 6).

In comparison to current bioactive peptide discovery methods, PepSAVI-MS boasts increased speed of identification by requiring minimal fractionation while maintaining a direct measure of activity. PepSAVI-MS does not require *a priori* knowledge of an organisms' genome or known biosynthetic clusters that produce bioactive peptides. Furthermore, there is no inherent bias towards highly active or abundant compounds. Together, these unique aspects of PepSAVI-MS allow for the high-throughput exploration of a suite of natural products with the opportunity to identify new and exciting classes of bioactive peptides.

1.7 Figures

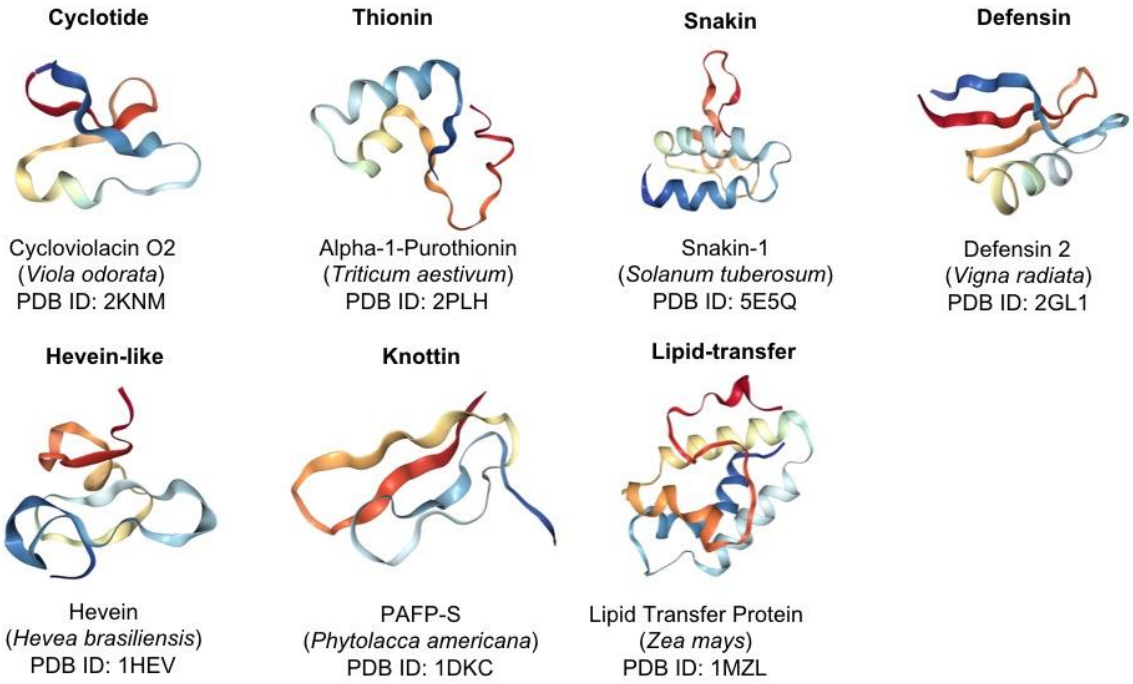


Figure 1.1 Plant antimicrobial peptide families. Representative structures gathered from the protein data bank using the NGL viewer[74-75].

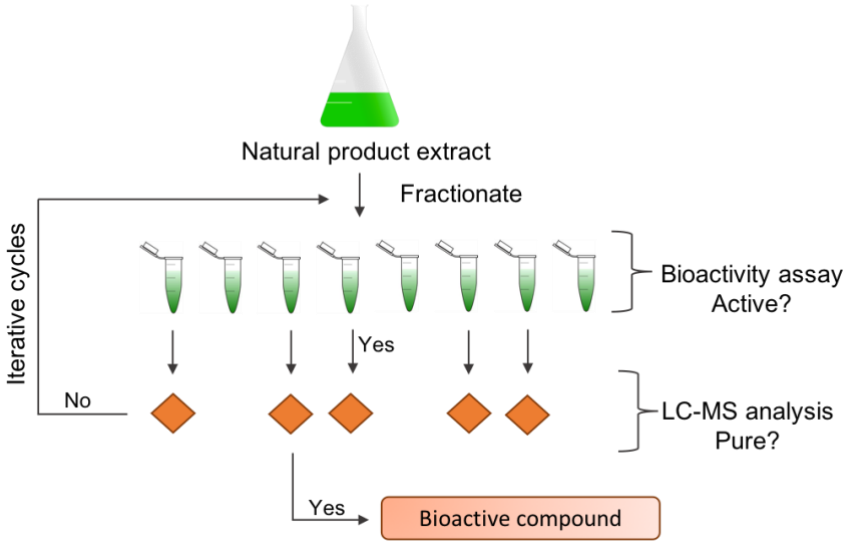


Figure 1.2 Bioassay guided fractionation schematic.

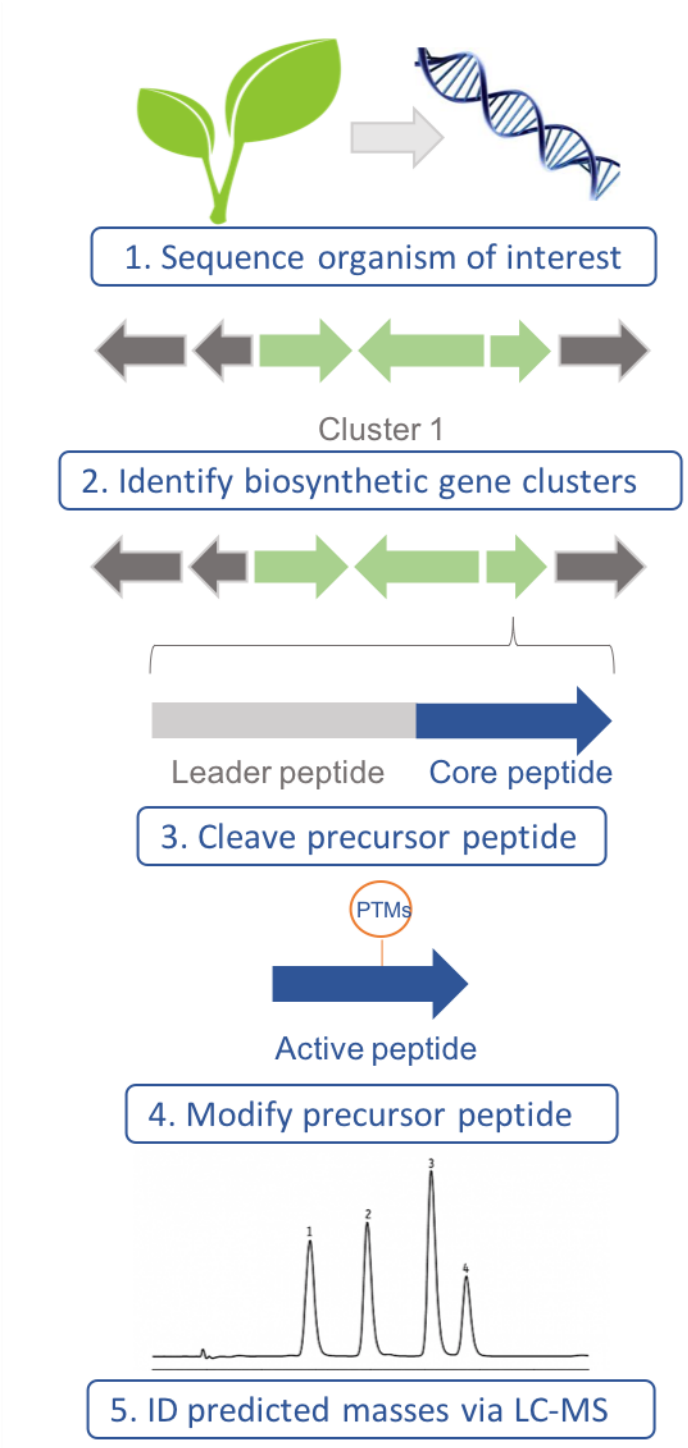


Figure 1.3 Genomic prediction schematic.

REFERENCES

1. O'Neill, J., Antimicrobial Resistance: tackling a crisis for the health and wealth of nations. The Review on Antimicrobial Resistance. **2014**.

2. Antibiotic resistance threats in the United States. Center for Disease Control and Prevention, **2013**.

3. Silver, L. L., Challenges of antibacterial discovery. *Clinical microbiology reviews* **2011,** *24* (1), 71-109.

4. Jones, D., The antibacterial lead discovery challenge. *Nature Reviews Drug Discovery* **2010,** *9* (10), 751-752.

5. Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L., Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* **2007,** *6* (1), 29-40.

6. Roemer, T.; Boone, C., Systems-level antimicrobial drug and drug synergy discovery. *Nature chemical biology* **2013,** *9* (4), 222-231.

7. Boucher, H. W.; Talbot, G. H.; Benjamin, D. K., Jr.; Bradley, J.; Guidos, R. J.; Jones, R. N.; Murray, B. E.; Bonomo, R. A.; Gilbert, D.; Infectious Diseases Society of, A., 10 x '20 Progress--development of new drugs active against gram-negative bacilli: an update from the Infectious Diseases Society of America. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **2013,** *56* (12), 1685-1694.

8. Boucher, H. W.; Talbot, G. H.; Bradley, J. S.; Edwards, J. E.; Gilbert, D.; Rice, L. B.; Scheld, M.; Spellberg, B.; Bartlett, J., Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **2009,** *48* (1), 1-12.

9. Klevens, R. M.; Edwards, J. R.; Richards, C. L., Jr.; Horan, T. C.; Gaynes, R. P.; Pollock, D. A.; Cardo, D. M., Estimating health care-associated infections and deaths in U.S. hospitals, 2002. *Public Health Rep.* **2007,** *122* (2), 160-166.

10. Antibiotic resistance threats in the United States. ***2013*** *[http://www.cdc.gov/drugresistance/threat-report-2013/]*.

11. Oerke, E. C., Crop losses to pests. *The Journal of Agricultural Science* **2006,** *144* (1), 31-43.

12. Figueroa, M.; Hammond-Kosack, K. E.; Solomon, P. S., A review of wheat diseases-a field perspective. *Mol. Plant Pathol.* **2017**.

13. Gu, F.; Khimani, A.; Rane, S. G.; Flurkey, W. H.; Bozarth, R. F.; Smith, T. J., Structure and function of a virally encoded fungal toxin from Ustilago maydis: a fungal and mammalian Ca2+ channel inhibitor. *Structure* **1995,** *3* (8), 805-814.

14. Appelbaum, P. C., 2012 and beyond: potential for the start of a second pre-antibiotic era? *The Journal of antimicrobial chemotherapy* **2012,** *67* (9), 2062-2068.

15. McGann, P.; Snesrud, E.; Maybank, R.; Corey, B.; Ong, A. C.; Clifford, R.; Hinkle, M.; Whitman, T.; Lesho, E.; Schaecher, K. E., Escherichia coli Harboring mcr-1 and blaCTX-M on a Novel IncF Plasmid: First report of mcr-1 in the USA. *Antimicrobial agents and chemotherapy* **2016,** *60* (7), 4420-4421.

16. Brown, E. D., Is the GAIN Act a turning point in new antibiotic discovery? *Canadian journal of microbiology* **2013,** *59* (3), 153-156.

17. Hughes, D.; Karlen, A., Discovery and preclinical development of new antibiotics. *Upsala journal of medical sciences* **2014,** *119* (2), 162-169.

18. Vlieghe, P.; Lisowski, V.; Martinez, J.; Khrestchatisky, M., Synthetic therapeutic peptides: science and market. *Drug discovery today* **2010,** *15* (1-2), 40-56.

19. WHO Model List of Essential Medicines - 20th List. **2017,** *http://www.who.int/medicines/publications/essentialmedicines/en/*.

20. Devocelle, M., Targeted antimicrobial peptides. *Frontiers in immunology* **2012,** *3* (309), 1-4.

21. Haney, E. F.; Nguyen, L. T.; Schibli, D. J.; Vogel, H. J., Design of a novel tryptophan-rich membrane-active antimicrobial peptide from the membrane-proximal region of the HIV glycoprotein, gp41. *Beilstein journal of organic chemistry* **2012,** *8*, 1172-1184.

22. Fjell, C. D.; Hiss, J. A.; Hancock, R. E.; Schneider, G., Designing antimicrobial peptides: form follows function. *Nature reviews. Drug discovery* **2012,** *11* (1), 37-51.

23. Mandal, S. M.; Roy, A.; Ghosh, A. K.; Hazra, T. K.; Basak, A.; Franco, O. L., Challenges and future prospects of antibiotic therapy: from peptides to phages utilization. *Front. Pharmacol.* **2014,** *5* (105), 1-12.

24. Scognamiglio, P. L.; Di Natale, C.; Perretta, G.; Marasco, D., From peptides to small molecules: an intriguing but intricated way to new drugs. *Current medicinal chemistry* **2013,** *20* (31), 3803-3817.

25. Walensky, L. D.; Bird, G. H., Hydrocarbon-Stapled Peptides: Principles, Practice, and Progress. *J Med Chem* **2014,** *57* (15), 6275-6288.

26. Verdine, G. L.; Hilinski, G. J., Stapled peptides for intracellular drug targets. *Methods in enzymology* **2012,** *503*, 3-33.

27. Wong, C. T.; Rowlands, D. K.; Wong, C. H.; Lo, T. W.; Nguyen, G. K.; Li, H. Y.; Tam, J. P., Orally active peptidic bradykinin B1 receptor antagonists engineered from a cyclotide scaffold for inflammatory pain treatment. *Angewandte Chemie* **2012,** *51* (23), 5620-5624.

28. O'Banion, C. P.; Nguyen, L. T.; Wang, Q.; Priestman, M. A.; Holly, S. P.; Parise, L. V.; Lawrence, D. S., The Plasma Membrane as a Reservoir, Protective Shield, and Light-Triggered Launch Pad for Peptide Therapeutics. *Angew. Chem. Int. Ed. Engl.* **2016,** *55* (3), 950-954.

29. Rezaei Araghi, R.; Bird, G. H.; Ryan, J. A.; Jenson, J. M.; Godes, M.; Pritz, J. R.; Grant, R. A.; Letai, A.; Walensky, L. D.; Keating, A. E., Iterative optimization yields Mcl-1–targeting stapled peptides with selective cytotoxicity to Mcl-1–dependent cancer cells. *Proceedings of the National Academy of Sciences* **2018,** *115* (5), E886-E895.

30. Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D., The future of peptide-based drugs. *Chem. Biol. Drug Des.* **2013,** *81* (1), 136-147.

31. Zlitni, S.; Ferruccio, L. F.; Brown, E. D., Metabolic suppression identifies new antibacterial inhibitors under nutrient limitation. *Nature chemical biology* **2013,** *9* (12), 796-804.

32. Ling, L. L.; Schneider, T.; Peoples, A. J.; Spoering, A. L.; Engels, I.; Conlon, B. P.; Mueller, A.; Schaberle, T. F.; Hughes, D. E.; Epstein, S.; Jones, M.; Lazarides, L.; Steadman, V. A.; Cohen, D. R.; Felix, C. R.; Fetterman, K. A.; Millett, W. P.; Nitti, A. G.; Zullo, A. M.; Chen, C.; Lewis, K., A new antibiotic kills pathogens without detectable resistance. *Nature* **2015,** *517* (7535), 455-459.

33. Kang, H. S.; Brady, S. F., Mining soil metagenomes to better understand the evolution of natural product structural diversity: pentangular polyphenols as a case study. *Journal of the American Chemical Society* **2014,** *136* (52), 18111-18119.

34. Milshteyn, A.; Schneider, J. S.; Brady, S. F., Mining the metabiome: identifying novel natural products from microbial communities. *Chemistry & biology* **2014,** *21* (9), 1211-1223.

35. Bachmann, B. O.; Van Lanen, S. G.; Baltz, R. H., Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *Journal of industrial microbiology & biotechnology* **2014,** *41* (2), 175-184.

36. Cichewicz, R. H., Epigenome manipulation as a pathway to new natural product scaffolds and their congeners. *Natural product reports* **2010,** *27* (1), 11-22.

37. Newman, D. J.; Cragg, G. M., Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J. Nat. Prod.* **2012,** *75* (3), 311-335.

38. Soejarto, D. D.; Fong, H. H.; Tan, G. T.; Zhang, H. J.; Ma, C. Y.; Franzblau, S. G.; Gyllenhaal, C.; Riley, M. C.; Kadushin, M. R.; Pezzuto, J. M.; Xuan, L. T.; Hiep, N. T.; Hung, N. V.; Vu, B. M.; Loc, P. K.; Dac, L. X.; Binh, L. T.; Chien, N. Q.; Hai, N. V.; Bich, T. Q.; Cuong, N. M.; Southavong, B.; Sydara, K.; Bouamanivong, S.; Ly, H. M.; Thuy, T. V.; Rose, W. C.; Dietzman, G. R., Ethnobotany/ethnopharmacology and mass

bioprospecting: issues on intellectual property and benefit-sharing. *Journal of ethnopharmacology* **2005,** *100* (1-2), 15-22.

39. McChesney, J. D.; Venkataraman, S. K.; Henri, J. T., Plant natural products: back to the future or into extinction? *Phytochemistry* **2007,** *68* (14), 2015-2022.

40. Gurib-Fakim, A., Medicinal plants: traditions of yesterday and drugs of tomorrow. *Molecular aspects of medicine* **2006,** *27* (1), 1-93.

41. Jones, J. D.; Dangl, J. L., The plant immune system. *Nature* **2006,** *444* (7117), 323-329.

42. Zasloff, M., Antimicrobial peptides of multicellular organisms. *Nature* **2002,** *415* (6870), 389-395.

43. Hancock, R. E.; Chapple, D. S., Peptide antibiotics. *Antimicrobial agents and chemotherapy* **1999,** *43* (6), 1317-1323.

44. Hancock, R. E.; Diamond, G., The role of cationic antimicrobial peptides in innate host defences. *Trends in microbiology* **2000,** *8* (9), 402-410.

45. Barbosa Pelegrini, P.; Del Sarto, R. P.; Silva, O. N.; Franco, O. L.; Grossi-de-Sa, M. F., Antibacterial peptides from plants: what they are and how they probably work. *Biochem Res Int* **2011,** *2011*, 250349.

46. Phoenix, D.; Dennison, S. R.; Harris, F., *Antimicrobial peptides*. Wiley Press. **2013**, 248 pages.

47. Pasupuleti, M.; Schmidtchen, A.; Malmsten, M., Antimicrobial peptides: key components of the innate immune system. *Critical reviews in biotechnology* **2012,** *32* (2), 143-171.

48. Goransson, U.; Burman, R.; Gunasekera, S.; Stromstedt, A. A.; Rosengren, K. J., Circular proteins from plants and fungi. *The Journal of biological chemistry* **2012,** *287* (32), 27001-27006.

49. Hammami, R.; Ben Hamida, J.; Vergoten, G.; Fliss, I., PhytAMP: a database dedicated to antimicrobial plant peptides. *Nucleic acids research* **2009,** *37* (Database issue), D963-968.

50. Wang, G.; Li, X.; Wang, Z., APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic acids research* **2016,** *44* (D1), D1087-1093.

51. Shim, Y. Y.; Young, L. W.; Arnison, P. G.; Gilding, E.; Reaney, M. J., Proposed systematic nomenclature for orbitides. *Journal of natural products* **2015,** *78* (4), 645-652.

52. Candido-Bacani Pde, M.; Figueiredo Pde, O.; Matos Mde, F.; Garcez, F. R.; Garcez, W. S., Cytotoxic Orbitide from the Latex of Croton urucurana. *Journal of natural products* **2015,** *78* (11), 2754-2760.

53. Segura, A.; Moreno, M.; Madueno, F.; Molina, A.; Garcia-Olmedo, F., Snakin-1, a peptide from potato that is active against plant pathogens. *Molecular plant-microbe interactions : MPMI* **1999,** *12* (1), 16-23.

54. Herbel, V.; Wink, M., Mode of action and membrane specificity of the antimicrobial peptide snakin-2. *PeerJ* **2016,** *4*, e1987.

55. Carvalho Ade, O.; Gomes, V. M., Plant defensins--prospects for the biological functions and biotechnological properties. *Peptides* **2009,** *30* (5), 1007-1020.

56. Ireland, D. C.; Colgrave, M. L.; Craik, D. J., A novel suite of cyclotides from Viola odorata: sequence variation and the implications for structure, function and stability. *The Biochemical journal* **2006,** *400* (1), 1-12.

57. Kommineni, S.; Bretl, D. J.; Lam, V.; Chakraborty, R.; Hayward, M.; Simpson, P.; Cao, Y.; Bousounis, P.; Kristich, C. J.; Salzman, N. H., Bacteriocin production augments niche competition by enterococci in the mammalian gastrointestinal tract. *Nature* **2015,** *526* (7575), 719-722.

58. Scherlach, K.; Graupner, K.; Hertweck, C., Molecular bacteria-fungi interactions: effects on environment, food, and medicine. *Annu. Rev. Microbiol.* **2013,** *67*, 375-397.

59. Dang, T.; Süssmuth, R. D., Bioactive Peptide Natural Products as Lead Structures for Medicinal Use. *Acc. Chem. Res.* **2017,** *50* (7), 1566-1576.

60. Essig, A.; Hofmann, D.; Munch, D.; Gayathri, S.; Kunzler, M.; Kallio, P. T.; Sahl, H. G.; Wider, G.; Schneider, T.; Aebi, M., Copsin, a novel peptide-based fungal antibiotic interfering with the peptidoglycan synthesis. *J. Biol. Chem.* **2014,** *289* (50), 34953-34964.

61. Bomgardner, M. M., Spider venom: An insecticide whose time has come? *Chem. Eng. News* **2017,** *95* (11), 30-31.

62. Golneshin, A.; Adetutu, E.; Ball, A. S.; May, B. K.; Van, T. T.; Smith, A. T., Complete Genome Sequence of Lactobacillus plantarum Strain B21, a Bacteriocin-Producing Strain Isolated from Vietnamese Fermented Sausage Nem Chua. *Genome Announc* **2015,** *3* (2), e00055-00015.

63. Harvey, A. L., Natural products in drug discovery. *Drug discovery today* **2008,** *13* (19-20), 894-901.

64. Anjum, T.; Bajwa, R., Isolation of bioactive allelochemicals from sunflower (variety Suncross-42) through fractionation-guided bioassays. *Natural product research* **2010,** *24* (18), 1783-1788.

65. Weller, M. G., A unifying review of bioassay-guided fractionation, effect-directed analysis and related techniques. *Sensors* **2012,** *12* (7), 9181-9209.

66. Stromstedt, A. A.; Felth, J.; Bohlin, L., Bioassays in natural product research - strategies and methods in the search for anti-inflammatory and antimicrobial activity. *Phytochem. Anal* **2014,** *25* (1), 13-28.

67. Molinari, G., Natural products in drug discovery: present status and perspectives. *Advances in experimental medicine and biology* **2009,** *655*, 13-27.

68. Skinnider, M. A.; Johnston, C. W.; Edgar, R. E.; Dejong, C. A.; Merwin, N. J.; Rees, P. N.; Magarvey, N. A., Genomic charting of ribosomally synthesized natural product chemical space facilitates targeted mining. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113* (42), E6343-e6351.

69. Medema, M. H.; Paalvast, Y.; Nguyen, D. D.; Melnik, A.; Dorrestein, P. C.; Takano, E.; Breitling, R., Pep2Path: automated mass spectrometry-guided genome mining of peptidic natural products. *PLoS Comput. Biol.* **2014,** *10* (9), e1003822.

70. Mohimani, H.; Kersten, R. D.; Liu, W. T.; Wang, M.; Purvine, S. O.; Wu, S.; Brewer, H. M.; Pasa-Tolic, L.; Bandeira, N.; Moore, B. S.; Pevzner, P. A.; Dorrestein, P. C., Automated genome mining of ribosomal peptide natural products. *ACS Chem. Biol.* **2014,** *9* (7), 1545-1551.

71. Mohimani, H.; Liu, W.-T.; Kersten, R. D.; Moore, B. S.; Dorrestein, P. C.; Pevzner, P. A., NRPquest: Coupling Mass Spectrometry and Genome Mining for Nonribosomal Peptide Discovery. *J. Nat. Prod.* **2014,** *77* (8), 1902-1909.

72. Potterat, O.; Hamburger, M., Natural products in drug discovery - Concepts and approaches for tracking bioactivity. *Curr Org Chem* **2006,** *10* (8), 899-920.

73. Queiroz, E. F.; Wolfender, J. L.; Hostettmann, K., Modern approaches in the search for new lead antiparasitic compounds from higher plants. *Current drug targets* **2009,** *10* (3), 202-211.

74. Rose, A. S.; Hildebrand, P. W., NGL Viewer: a web application for molecular visualization. *Nucleic Acids Res.* **2015,** *43* (W1), W576-W579.

75. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res.* **2000,** *28* (1), 235-242.

CHAPTER 2: Detailed PepSAVI-MS Protocols

2.1 Introduction

PepSAVI-MS is a streamlined pipeline to facilitate natural product bioactive peptide discovery and consists of the following major steps: plant or microbial growth and harvesting, peptide extraction, peptide library generation, bioactivity screening, LC-MS/MS, and statistical analysis to identify bioactive components. While previous experiments in the Hicks Laboratory influenced many of the techniques and procedures described in this chapter, all of the protocols required modification for use in PepSAVI-MS. As such, this chapter serves as detailed reference manual for the techniques laying the foundation of PepSAVI-MS. The protocols discussed below serve as a starting point for new experiments, but may require adjustments or optimization for each individual species, bioactivity assay, or available LC-MS platform.

2.2 Plant Growth and Harvesting

Plant species are selected based on criteria including known antimicrobial bioactivity, traditional medicinal properties, and availability. Seeds or seedlings are obtained through commercial sources (e.g. Horizon Herbs, Southern Exposure Seed Exchange) that carry out rigorous testing to ensure correct identity of seeds, plants, and roots. Upon arrival, seeds are stored at room temperature in original packaging until planted in the UNC Greenhouse Facility. The greenhouse is kept between 63.5 and 68.5°F and is on a 14-hour daylight cycle. Seeds of each species are planted in the appropriate soil for optimal growth conditions and are allowed to grow until ample tissue is available (~6-10 weeks). For species that are difficult to germinate, a variety of modifications/alternate growth options have been explored. The fastest and easiest

solution is to order the live plant, if available, and transplant into standard greenhouse soil/conditions. If this option is used (e.g. *Viola odorata*), the live plant should be kept alive as long as possible to ensure sufficient collection of aerial tissue. Other methods to increase germination include shocking or scratching the seed coat, addition of gibberellic acid to the soil, or growth in a greenhouse chamber with optimized temperature and humidity conditions (see the notebook "Danny Lee 2016-2017/Wilaysha Evans" for specific growth conditions for each species). These methods have been used with varying degrees of success and require research of optimal growth conditions of each individual species before use. Regardless of growth strategy, healthy aerial tissue, excluding stem mass if possible, is transferred to aluminum packets, flash frozen under liquid nitrogen, and stored at -80°C to preserve protein structure and function.

2.3 Peptide Extraction

Nearly all known AMPs have four important and relatively conserved physical properties: they are cysteine rich, highly basic, amphipathic, and <10 kDa in size[1]. Current extraction protocols exploit these properties to selectively extract for bioactive peptide-like molecules: 1) aqueous buffers are used to selectively extract highly basic compounds and 2) size exclusion steps are used to select for the mass range of known bioactive peptides. Extraction protocols for two classes of natural products (plants and microbes) have been developed. All methods for peptide extraction result in the creation of a crude extract, which is used as the starting material for peptide library creation.

2.3.1 Plants

Frozen plant tissue is ground and extracted into aqueous buffer. Size exclusion steps via dialysis and molecular weight cutoff filters are used to select for the mass range of bioactive peptides. Extracts are concentrated using vacuum centrifugation and concentrated to a common

volume for preparation of crude extracts, which are stored at -80°C until further processing. The full protocol takes ~4 days to complete and is detailed as follows:

Day 1:

1. Estimate the tissue mass of the extraction species. Optimal extraction conditions use 100 g of ground starting material and volumes listed below correspond to this weight. If less tissue is available, scale all reagents accordingly.
2. Prepare 300 mL of acetic acid extraction buffer using the following recipe:
    a. 10% acetic acid (30 mL), 90% water (270 mL).
    b. Add 6 Roche EDTA-free protease inhibitor cocktail tablets.
    c. Add pepstatin (0.5 mg/mL stock, in 80% ethanol) to a final concentration of 1 mg/L.
    d. Stir solution containing acetic acid, water, and protease inhibitors until plant tissue is ground.
    e. Right before addition of ground material, add 3% polyvinylpolypyrrolidone (PVPP) (g/mL). PVPP is insoluble in aqueous solution and will make the solution appear milky.
3. Pre-chill the mortar in a nest of dry ice, covered in aluminum foil. Keep mortar in dry ice while grinding tissue.
4. Add liquid nitrogen (N_2(l)) to empty mortar, add batch of plant tissue and add more N_2(l).
5. Grind tissue until pulverized to a fine powder. Add N_2 (l) as needed so that plant material stays frozen throughout the grinding process.

6. Allow liquid nitrogen to boil off and add pulverized batch to weigh boat. Continue to grind until this weight reaches 100 g. If storing ground tissue in the weigh boat while continuing to grind, be sure to store ground tissue at -80℃.
7. Once 100 g ground material is accumulated, immediately add the PVPP and then plant material to the beaker with extraction buffer and stir for 15 minutes at room temperature.
8. Transfer beaker to 4℃ and stir for 4 h.
9. Transfer extract into centrifuge bottles and centrifuge for 45 minutes at 13,200 RCF, 4°C.
10. Gently remove bottles after centrifugation to avoid disturbing the insoluble pellet and pour supernatant into the top portion of a Stericup (Millipore, 0.45 μm, HV Durapore Membrane).
11. Vacuum filter the liquid material through the Stericup. This process is highly variable depending on the plant species – can take anywhere from 5 minutes to multiple days to pull material through. Keep extract cold and on ice if this process is on the day(s) timescale.
12. Store the filtered extract at 4℃ overnight.

Day 2:

13. Prepare Amicon Ultra-15 Centrifugal Filter Devices, 30000 MWCO (Millipore, UFC9010024) per manufacturer protocol to remove glycerin. Each filter holds 15 mL of extract and can be used twice for the same extract (estimate volume in Stericup and calculate accordingly).
14. Split material equally among Amicon Ultra-15 tubes in 15 mL portions.

15. Centrifuge for 1 hour, at 3,220 RCF, 4°C. Check to see if the majority of the extract has gone through the filter. If so, remove and save the bottom portion and add another round of extract to the top of the filter and repeat. If not, continue with the centrifugation until most of the material is pulled through and then repeat.
16. Recombine the filtrate into one container and save high-MW material for later analysis (store at -80°C).
17. Pipette 5 mL of filtrate into 8 mL plastic tubes until all filtrate is used. Freeze at -20℃.

Day 2 Overnight into Day 3:

18. Once frozen, concentrate the material via vacuum centrifugation with no heat for 4-12 hours depending on volume of extract. Do not concentrate to dryness, as this may lead to precipitation.
19. At completion, note the volume. Expect a 10-fold concentration (~30 mL remaining).

Day 3:

20. Prepare SpectraPor Float-A-Lyzer G2 500-1000 MWCO 10 mL devices (G235051 according to manufacturer instruction (wash with dilute ethanol, then water).
21. During the conditioning step, prepare one 4 L Nalgene bucket of 5 mM ammonium formate (HPLC or MS Grade; Fluka 17843/70221).
22. Adjust the buffer pH to 2.7 using formic acid.
23. Dialyze concentrated extract into 5 mM ammonium formate, pH 2.7 using four preconditioned devices (~10 mL/device).
24. Perform two dialysis steps, changing out the buffer in between; first dialysis for 4-6 h, then overnight for second. Dialyze with stirring at 4°C.

Day 4:

25. Carefully remove the dialysis tubes from the buckets and draw out remaining liquid using long plastic dropper pipettes.
26. Put approximately 5 mL dialyzed material into 8 mL plastic tubes and freeze at -20℃.
27. Vacuum concentrate material following dialysis, with no heat, for 4-12 hours depending on volume of extract. Pool material into 2 mL micro centrifuge tubes as the process progresses. Do not concentrate to dryness, as this may lead to precipitation.
28. Concentrate extract from 100 g of starting material (300 mL extraction buffer) down to 4 mL final volume. This concentration is what we are calling the "1X" concentration.
29. Aliquot into 420 µL portions and store at -80℃ until further use.

2.3.2 Microbial Secretomes

Modifications to the extraction protocol are required to capture microbial bioactive peptides, which are secreted into the environment for defense and niche security. As such, large-scale microbial cultures are grown at optimal conditions and cation exchange resin is use to capture secreted peptides from the cell-free media following the detailed protocols:

Fungi:

Day 1:

1. Prepare four 5 mL starter cultures of the microbe of interest and grow at optimal temperature (usually 25℃) with shaking.

Day 3:

2. Prepare 1 L each of 25 mM sodium acetate, pH 5.5 (equilibration/wash) and 25 mM sodium acetate, pH 5.5, with 1 M NaCl (elution) solutions.
3. Equilibrate CM Sephadex C-25 resin (GE Healthcare) by weighing out 10 g and hydrating in 140 mL of equilibration buffer. Stir slowly at 4℃ for two days.

Day 5:

4. Add all 4 starter cultures to 2 L of media and grow to late-log phase at optimal growth conditions.
5. Remove cells by centrifugation (500 x g for 5 minutes) and carefully decant supernatant into a new flask without disturbing the cell pellet.
6. Adjust supernatant to a pH of 5.5 and stir overnight with the hydrated resin.

Day 6:

7. Transfer peptide-resin slurry into centrifuge bottles and gently centrifuge (500 x g for 5 minutes) the beads to the bottom of the bottle.
8. Decant (save, and label as flow through) remaining liquid from the bottle. Add enough wash solution to cover the resin (~50 mL), gently swirl in bottles for ~5 minutes, and repeat the centrifugation process.
9. Continue to wash and collect resin 5x (save, label as wash #).
10. After 5 washes, add enough elution solution (~50 mL) to cover the resin, swirl, centrifuge, and decant (save, label as elution #). Repeat elution step 5 times.
11. Buffer exchange and concentrate elution fractions 10X into 5 mM ammonium formate, pH 2.7 using 3 kDa spin concentration filters (Millipore). Combine concentrated elutions for creation of crude extract.

Bacteria:

Day 1:

1. Prepare 1 L each of 25 mM sodium acetate, pH 5.5 (equilibration/wash) and 25 mM sodium acetate, pH 5.5, with 1 M NaCl (elution) solutions.
2. Equilibrate CM Sephadex C-25 resin (GE Healthcare) by weighing out 10 g and hydrating in 140 mL of equilibration buffer. Stir slowly at 4℃ for two days.

Day 2:

3. Prepare four 5 mL starter cultures of the microbe of interest and grow until late-log phase at optimal temperature (usually 37℃) with shaking.

Day 3:

4. Add all 4 starter cultures to 2 L of media and grow to late-log phase at optimal growth conditions.
5. Remove cells by centrifugation (500 x g for 5 minutes) and carefully decant supernatant into a new flask without disturbing the cell pellet.
6. Adjust supernatant to a pH of 5.5 and stir overnight with the hydrated resin.

Day 4:

7. Transfer peptide-resin slurry into centrifuge bottles and gently centrifuge (500 x g for 5 minutes) the beads to the bottom of the bottle.
8. Decant (save, and label as flow through) remaining liquid from the bottle. Add enough wash solution to cover the resin (~50 mL), gently swirl in bottles for ~5 minutes, and repeat the centrifugation process.

9. Continue to wash and collect resin 5x (save, label as wash #).
10. After 5 washes, add enough elution solution (~50 mL) to cover the resin, swirl, centrifuge, and decant (save, label as elution #). Repeat elution step 5 times.
11. Buffer exchange and concentrate elution fractions 10X into 5 mM ammonium formate, pH 2.7 using 3 kDa spin concentration filters (Millipore). Combine concentrated elutions for creation of crude extract.

2.4 Peptide Library Generation

After selective extraction (plants) or harvesting (microbes) of small cationic peptides, extracts are crudely fractionated using strong cation exchange chromatography on a PolySulfoethylA column (100 x 4.6 mm, 3 μm particles). Briefly, 400 μL of each extract is injected on column and eluted using a high salt gradient (mobile phase A: 5 mM ammonium formate, 20% ACN, pH 2.7; mobile phase B: 500 mM ammonium formate, 20% ACN, pH 3.0). Salt gradients have been optimized for plant extracts (Figure 2.1a) and microbial extracts (Figure 2.1b). Fractions are collected in one-minute intervals across the entire run and desalted 3x with MilliQ water using a vacuum concentrator to remove salts toxic to microbial cells. Desalted fractions are stored for short periods at 4°C if downstream analysis will happen immediately after desalting, or at -20℃ until further use. Freeze-thaw cycles should be minimized and bioactivity screening and LC-MS analysis should happen as close together as possible. A detailed protocol for Shimadzu HPLC operation, sample injection, and fraction collection in the context of peptide library creation has been developed and is described below:

1. Preparation of Mobile Phases:
    a. Prepare appropriate mobile phases in 1L glass bottles.

b. Sonicate mobile phases in a bath sonicator with loosened caps for 1 hour to de-gas solutions.

c. Save ~10 mL of MPA in a falcon tube to use as a blank.

d. With the pumps turned off, switch out the mobile phases by removing the cap and attached tubing off of the first bottle, rinsing with dH_2O, wiping with a kimwipe, and inserting into new mobile phase. Repeat with mobile phase B. (Note: The lines are labeled with A and B to denote which mobile phase is which).

2. Sample Preparation:

   a. Thaw extract and centrifuge for 5 minutes at 5000 rpm to pellet out any precipitation.

   b. Add sample to 1 mL vial with at 10% more than the injection volume. Make sure there are no air bubbles in the vial.

3. Changing the column:

   a. Unscrew the fittings on the old column, undoing the top ferrule first.

   b. Screw in the bottom ferrule of the new column, then the top ferrule. Fittings should be finger tight. The flow goes from the bottom of the chamber to the top, so make sure the flow arrow is pointing upwards.

   c. Turn on the pump and watch the column to make sure there are no leaks.

4. Creating a Method:

   a. Ensure all components of the LC are turned on and display a green light.

   b. Click on "Instruments."

   c. Double click on "System 1" – should hear a single beep to know instrument was connected properly and you should see "LC Ready" in the upper left corner.

   d. Open a new method file: File-> New method file.

e. Expand the Instrument parameters view section and navigate the following tabs:

i. Data Acquisition:

(a) Change “LC stop time” and “End time” to your run duration.

ii. LC Time Program:

(a) Write gradient program.

(b) Press “Draw curve” to ensure your gradient is written correctly.

iii. Pump:

(a) Set “Flow” to your desired flow rate.

(b) Change the maximum pressure limit to 4500 psi and the minimum pressure limit to 0 psi.

iv. Detector A:

(a) Use the default settings, changing only the desired detection wavelengths.

v. Column Oven:

(a) Set temperature if desired.

vi. Controller:

(a) Under external output, check power on.

(b) Turn autosampler pretreatment off.

vii. Autosampler:

(a) Use default settings.

viii. Autopurge:

(a) 1st: Mobile Phase A for 10 minutes.

(b) 2nd: Mobile Phase B for 10 minutes.

(c) Check Autosampler box, for 25 minutes.

(d) Under Warmup:

(i) Wait time: 0 minutes.

(ii) Total flow: 0.5 mL/minutes.

(e) Check activate system after autopurge.

f. Save your method file.

5. Starting the Instrument:

a. "Wake up" all hardware manually by pressing the sleep button. Turn on the detector by pressing the power button (Orange=Asleep, Green=Awake, Red=Problem).

b. User Name: LC User, Password: LC User.

c. Double click "LabSolutions" icon.

d. Click on "Instruments."

e. Double click on "System 1" – should here a single beep to know instrument was connected properly and you should see "LC Ready" in the upper left corner.

f. Open your method file and start the pump: File→Open.

g. Autopurge following the method settings. Note: You should never autopurge with the column oven on. If the column oven is turned on for your method, press the button for "Oven On/Off" before purging.

h. Run at least 1 blank (MPA) using your method to ensure the instrument is performing properly.

i. Start samples using either the single run, or by creating a batch upload.

i. Running a Sample: Single Run.

(a) Open the method file you wish to use.

(b) Press the single run button (green diamond).

(c) Name your Sample.

(d) Change the file where your data will be stored by clicking on the folder icon to the right of "Data file". Choose your folder and name the file the same as you named your sample.

(e) Adjust the vial position and injection volume accordingly (max 500 uL or 10 uL depending on installed sample loop). The tray should always be 1.

(f) Press "OK" to start the run.

ii. Running a Sample: Batch Upload.

(a) Open method.

(b) Click main in the left tab.

(c) Press "Realtime batch."

(d) Queue up samples.

(e) Press start realtime batch.

6. Fraction collection:

a. The fraction collector runs independently of the Shimadzu. If fraction collection is desired, it must be manually started at the beginning of each run.

b. Before starting the run, align the fraction collector by lining up the dashed line on the turn barrel with the dashed line on the stand. Ensure correct alignment by placing an empty tube in position 1 and make sure that the fraction collector is dripping into the center of the tube. The fraction collector or cannot be adjusted once it is started so ensure that alignment is accurate.

c. Place numbered tubes in the fraction collector in their correct positions.

d. Set how frequently the fraction collector should turn by adjusting the buttons on the front of the stand (turn after X number of drops or after X time passes).

e. Start the run on the Shimadzu and watch the computer. Once the sample is injected, the run time will officially begin and the fraction collector should be started such that the run time is representative of the fraction number.

7. Shutting down the instrument:

   a. Turn the pump off by pressing the "Pump On/Off" button in LabSolutions. Watch to make sure both the flow rate and pressure drop to zero on the instrument display. Note: The pump continues to pump 100% MPA after the run(s) have ended. It will run overnight and use all your mobile phase/suck up air if you forget to turn it off.

   b. Turn the column oven off by pressing "Oven On/Off" if applicable.

   c. Close both your method and LabSolutions. You should hear a beep to let you know the instrument is disconnected.

   d. Sleep the computer.

   e. Manually press the sleep button on both pumps, the autosampler, and the column oven (status lights should turn orange). Turn the detector completely off by pressing the power button until it pops out – this is to save lamp hours (status light should be off). The status light on the control bus should remain green.

8. Viewing Previous Runs:

   a. Open LabSolutions.

   b. Click Postrun on left.

   c. Double click Postrun icon.

   d. Open appropriate data folder.

e. Highlight, drag, and drop your desired file in top graph. Note: multiple files can be highlighted at once and dragged/dropped together to compare two or more runs.
f. Click and drag around trace on top graph to zoom in on features. The zoomed in section will show up on the bottom graph.

9. Troubleshooting:
    a. Refer to the troubleshooting guide on the Hicks Lab Sharedrive (Lab_Documents/Protocols/Shimadzu HPLC Troubleshooting Guide) for common issues.

2.5 Bioactivity Screening

Bioactive peptides have been shown to exhibit a wide range of bioactivities, including antibacterial, antifungal, and anticancer activity. As such, assays have been developed to probe the intrinsic bioactivities of natural product peptide libraries generated with PepSAVI-MS. All assays rely on the addition of the peptide library to an early log phase microbial culture or cell line, and subsequent quantification of remaining viable cells after an optimized incubation period. For all assays, it is crucial to use correct sterile technique for accurate and reproducible results.

2.5.1 Antibacterial Activity

For initial antibacterial screening, an assay against the gram-negative bacterium *Escherichia coli* (ATCC 25922) has been developed. Peptide libraries are pre-screened for strong activity against *E. coli* to prioritize fraction sets for further screening against the clinically relevant ESKAPE pathogens. Bioactivity assays for a subset of the ESKAPE pathogens are currently being established and optimized in our laboratory. A collaboration with the laboratory of Dr. Les Shaw at the University of South Florida has been established for screening against the full panel

of ESKAPE pathogens. For any bacterial pathogen, cell culture is mixed with each plant fraction in microtiter plates and is incubated for 2-4 hours to allow sufficient time for bacterial growth. Cell viability is assessed using Resazurin, an oxidation-reduction indicator dye that quantifies living bacterial cells[2]. Percent activity is calculated using the formula: $\%\ Activity = \left(1 - \left(\frac{RFU(fraction) - RFU(positive\ control)}{RFU(negative\ control) - RFU(postive\ control)}\right)\right) * 100\%$. The detailed protocol for antibacterial assay setup and analysis is as follows:

Day 1:

1. Streak a Mueller Hinton agar plate with bacteria from the appropriate freezer stock. Use quadrant streaking to produce isolated colonies (16 hours at 37°C for *E. coli*). Store at 4°C between uses. Plates are stable at 4℃ for one week.

Day 2:

2. Using an inoculating stick, transfer a single colony from the plate to 5 ml of Mueller Hinton Broth (MHB, Difco) in a 14 mL culture tube. Incubate at 37°C with shaking (250-325 rpm) overnight (16 hour incubation e.g. 5pm – 9 am or 6 pm – 10 am).

Day 3:

3. Check optical density at 600 nm and subculture appropriately to produce (after further growth for 1 hour at 37°C) a culture with an OD_{600} of 0.5. If the OD is above 1.0, dilute your culture to obtain an accurate read.
    a. For *S. aureus* and *A. baumannii*, calculate volume of culture needed for final density 0.35 in 5 ml MHB.
    b. For *E. coli* and *K. pneumoniae*, calculate volume of culture needed for final density 0.25 in 5 ml MHB.

c. Use equation: $M_1V_1 = M_2V_2$

i. i.e. volume to add = $[(5000)(0.25 \text{ or } 0.35)]/OD_{600}$

4. After 1 hour additional growth, confirm culture density is ~0.5.
5. Refer to the bioassay plate layout template on the Hicks Lab Sharedrive (Lab_Documents/Antimicrobial Peptides/Bioassay Formatting Guidelines_revised_2) to determine the location of each component of the assay. Refrain from using the outermost wells of the plate to prevent edge effects. Add 50 µL of water to these wells to help prevent drying.
6. For one assay, performed in triplicate, make an 8 ml "master mix" by combining 2 mL of 2x MHB (Difco), 4 mL of 1x MHB (Difco), and 2 mL of culture. 2x MHB added to master mix will always be constant but 1x MHB and culture volumes can be adjusted to give correct final OD of 0.125 (This will be 0.10 when fractions are added).
7. Using a sterile 96 well microtiter dish, a sterile pipetting reservoir, and a multichannel pipette, add 40 µL master mix to appropriate wells.
8. For every replicate, include 4 wells of positive control using 10 µL of antibiotic.
    a. *E. coli* – use 10 µL of a 1:100 dilution of Amp_{100} stock (ampicillin, dilute in water).
    b. *S. aureus, A. baumannii, K. pneumoniae,* and *E. faecium* – use 10 µL of a 1:100 dilution of Erm_{100} stock (erythromycin, dilute in water).
    c. Include 4 wells of positive control culture on each plate.
9. For every replicate, include 4 wells of a negative control using 10 µL of water in place of the test solution.

10. For every experiment include 4 blank wells with 50 µL of 1x media in place of the bacterial culture.
11. Handle the plate carefully and keep level to avoid splashing cultures among adjacent wells. Keep the lid on the plate at all times except when checking OD.
12. Spin plate briefly in centrifuge to remove bubbles and ensure bottom of each well is fully coated with culture/test solutions.
    a. Press and hold the "Short" button until the condensation is knocked off the lid (~1500 rpm).
13. Read OD_{600} of the plate and verify that the starting OD is about 0.1.
14. Transfer plate to 37℃ incubator and insert into plate holder. Grow to mid log-phase (OD_{600} of 0.5) with vigorous shaking (250 rpm) for ~4 hours. Resazurin produces reliable estimates of metabolic activity at any point in the log-phase, but mid log-phase is a convenient point.
15. Read OD_{600} of the plate to verify the growth phase of the culture. Use a multichannel pipette and a sterile pipetting reservoir to add 1 µL of 50 mM Resazurin (dissolved in water and filter sterilized at 0.22 micron) to every well. Centrifuge as in step 11a to ensure all resazurin and culture is in the bottom of each well.
16. Incubate at 37℃ for 1 hour with shaking.
17. Place the adapter/spacer in the plate reader and set parameters for a top read. Read fluorescence with excitation at 544 nm and emission at 590 nm.
18. Export data in .csv format. Use appropriate template file to copy and paste data for final calculations.

2.5.2 Antifungal Activity

While AMPs are reported to have broad range activity, the most commonly identified activity of plant bioactive peptides is against fungal pathogens[1]. Two protocols for antifungal activity have been developed, the first a high-throughput screening assay using liquid culture as described for the antibacterial assays, and the second an agar diffusion-based assay for fungal species that fail to grow to high densities in liquid culture, form fungal mats, or do not sporulate under these growth conditions. For liquid culture-based assays, the agriculturally relevant filamentous fungus *F. graminearum* was selected. Conidia spores from *F. graminearum* are used for this assay and as such, the following protocol was adapted from The Fusarium Lab Manual[3] for spore germination and collection:

1. Grow a fresh agar plate of *F. graminearium* from a frozen spore stock or punch of a preserved plate.
2. Grow fungus until a dense mat is present (~5 days, pink color should be present).
3. Punch fresh fungal disks from the mat using the back of a 1 mL pipette tip. You will need 3 disks/50 mL of media.
4. Make and autoclave 6 x 50 mL flasks of CMC media. Cool media in water bath.
5. While media is cooling, isolate 3 punches in an empty Petri dish.
6. Add 3 mL of water to the fungal plugs and cut up the plugs into small chunks with a razor blade.
7. Add the water/spore mixture to one of the 6 flasks.
8. Repeat steps 5-7 until all 6 flasks contain spores.
9. Grow flasks at 30℃ with shaking at 120 rpm for three days.
10. On the fourth day, remove flasks from incubator and vortex to break up spores.

11. Place a funnel with a layer of miracloth on top of a clean, autoclaved 1L flask.
12. Pour each 50 mL flask into the bigger flask through the miracloth-covered funnel.
13. Discard miracloth into autoclave trash and split filtered spores into 50 mL falcon tubes.
14. Centrifuge falcon tubes at 4000 x g for 20 minutes at room temperature.
15. Pour off media, resuspend spores in enough water to cover the pellet and spin at 4000 x g for 20 minutes at room temperature.
16. Pour off the water.
17. Add 25 mL of 20% glycerol to one tube.
18. Use this liquid to clean out and combine the spores into one tube. Add more glycerol if necessary.
19. Count the cells using the hemocytometer.
20. Dilute cells to 2.0 x 10^6 conidia/mL with 20% glycerol.
21. Prepare 2 mL aliquots from the dilute stock solution.
22. Flash freeze stocks and store at -80℃ until further use.

Thawed spore stocks are mixed and incubated with each plant fraction in microtiter plates following the adapted bacterial protocol (note: refer to The Fusarium Lab Manual[3] for media recipes):

1. For one assay, performed in triplicate, make an 8 ml "master mix" by combining 2 ml of Fusarium spores with 6 ml of minimal media (MM).
2. Using a sterile 96 well microtiter dish, a sterile pipetting reservoir, and a multichannel pipette, add 40 µL master mix to appropriate wells.
3. For every replicate, include 4 wells of positive control using 10 µL of antibiotic.
    a. *F. graminearum* - use 10 µL of 50 µg/mL Hygromycin B.

4. For every replicate, include 4 wells of a negative control using 10 μL of water.
5. For every replicate include 4 blank wells with 50 μL of MM in place of the bacterial culture.
6. Handle the plate carefully and keep level to avoid splashing cultures among adjacent wells. Keep the lid on the plate at all times except when checking OD.
7. Spin plate briefly in centrifuge to remove bubbles and ensure bottom of each well is fully coated with culture/test solutions.
    a. Use "Short" button on the centrifuge to spin plates until the condensation is knocked off the lid (~1500 rpm).
8. Read OD_{600} of the plate and verify that the starting OD is about 0.1.
9. Transfer plate to 25℃ incubator and insert into plate holder. Grow to mid log-phase (OD_{600} of 0.5) with vigorous shaking (300 rpm) for 2 days.
10. After 48 hours, read OD_{600} of the plate to verify the growth phase of the culture.
11. Add 10 μL of BactiterGlo to each well and return to incubator to shake for 5 minutes.
12. Read luminescence of each plate on the spectrophotometer using a 500 millisecond integration time.

The number of remaining viable fungal cells are determined in a luminescence assay using BacTiter-Glo (Promega) through the quantification of ATP released from the cell lysis. Free ATP reacts with luciferase present in the reaction mixture and generates light that is measurable via luminescence. Percent activity of each well is calculated as previously described for the antibacterial assays.

Antifungal bioactivity has also been established an agar diffusion-based assay. While this method is more labor intensive, it is especially useful for fungal species that fail to grow to high

densities in low-volume culture or are prone to the formation of fungal mats at high concentrations. For these agar diffusion assays, the challenge microbial species is infused into the agar. Wells are then carved into solidified agar plates and the peptide library is added, one fraction per well. Agar is prepared as a soft agar to allow for compound diffusion. After incubation, zones of inhibition form around wells containing active compounds. Measurement of each radius of inhibition allows for a measurement of activity across all library fractions. The detailed protocol for agar-based diffusion assays is described below:

1. Grow dense fungal starter cultures according to optimal growth conditions (typically 5 days, 25℃, 250 rpm).
2. On the 6th day, prepare UM media (or other appropriate media) according to manufacturer/recipe.
3. Add 1% agar to liquid media (2.5 g agar for 250 mL media) to make a "soft" agar.
4. Boil media to dissolve agar and autoclave.
5. Cool media to 48℃ in a water bath.
6. Add fungal cultures to cooled media in the ratio of 5 mL fungus/250 mL agar.
7. Pour fungus-infused agar into Petri dishes and let solidify at room temperature for 1 hour.
8. Once solidified, carve 4 wells/plate into the agar using the back of a 10 µL pipette tip.
    a. Use a sterile wooden stick to pick up each tip, carefully avoiding touching the top of the tip with your gloved hands.
    b. Punch all 4 wells with the same tip and then discard. Change pipette tip after each plate to avoid contamination.
    c. Once wells are carved, use a sterile wooden stick to scoop agar out of each well. Change wooden sticks in between each well.

9. Let plates sit for 30 minutes at room temperature before compound addition to dry the holes.
10. Add 50 μL of each compound to a single well in each plate. Fractions should be assessed for activity minimally in duplicate.
11. Use appropriate controls for each challenge species.
12. Move plates carefully into the static incubator at the appropriate temperature. Keep right-side-up until you are sure that all of the liquid has diffused into the agar (usually takes a couple of hours). Then, flip plates over so that condensation will form on the lid of the plate and not the agar.
13. Let plates grow for 3-5 days until measurable zones of inhibition are observed in the positive control.
14. Measure and record each radius of inhibition by measuring the diameter of inhibition and dividing by two.

While most often used to assess antifungal activity, these agar diffusion-based assays can also be used for recalcitrant bacterial species. This was the case with *Enterococcus faecalis* and as such, agar diffusion assays with minor modifications were used to assess the activity of this library as described in Chapter 5.

2.5.3 Anticancer Activity

Anticancer assays have been established in collaboration with Dr. David Hoskin at Dalhousie University in Nova Scotia and with Dr. Mike Emanuele at the University of North Carolina at Chapel Hill. Assays have been developed for preliminary screening against three cell lines: MDA-MB-231 breast cancer, OVCAR-3 ovarian cancer, and PC-3 prostate cancer. Screening against a noncancerous human dermal fibroblasts cell line has also been established for toxicity

studies. All of the anticancer screens performed in this dissertation are in collaboration with the Hoskin Lab. However, a newly established collaboration with the Emanuele Lab allows for easier and faster anticancer analysis, and will serve as a route for preliminary screening to identify highly active peptide libraries that will undergo the full anticancer panel.

Hoskin Lab Protocol:

Day 1:

1. Grow cells in the appropriate growth medium and plate into a 96-well flat-bottom microtiter plates at 2×10^4 cells/well.
2. Culture cells for 24 hours in a 37℃ humidified incubator to allow for formation of adherent monolayers.

Day 2:

3. Remove medium and replace with fresh serum-free medium with or without individual plant fractions to be assayed for cytotoxic activity.
4. Culture cells for an additional 24 hours.

Day 3:

5. Two hours before the end of culture, add 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) solution to each well at a final concentration of 0.5 mg/mL to measure mitochondrial succinate dehydrogenase activity.
6. Discard culture supernatants.
7. Solubilize formazan crystals in 0.1 mL of dimethyl sulfoxide.

8. Measure absorbance at 570 nm and compare to the control to determine percent reduction in viability, as reflected by the change in mitochondrial succinate dehydrogenase activity.

Emanuele lab protocol:

Day 1:

1. Seed 10,000 MDA-MB-231 breast cancer cells/100 μL media/well into a 96-well plate (93 wells with cells, 3 wells without cells).
2. Allow for overnight adherence.

Day 2:

3. In the morning, carefully remove media from each well without disturbing cells.
4. Replace with 90 μL of fresh media in each well.
5. Add 10 μL of each fraction to the cells. Initial activity assessment is performed in singlet. Any active peptide libraries will be re-screened in triplicate.
6. Incubate at 37℃ for 24 hours.

Day 3:

7. In the morning, carefully remove the media from each well without disturbing the cells.
8. Replace with 90 μL of fresh media in each well.
9. Prepare 5 mg/mL of MTT solution in PBS.
10. Add 10 μL of 5 mg/mL MTT solution to each well and incubate for 2 hours at 37℃.
11. Carefully remove media from each well without disturbing cells.
12. Add 100 μL of DMSO to each well and mix by pipetting up and down. Be sure to add 100 μL of DMSO to wells without cells as a blank.

13. Incubate at 37℃ for 15 minutes.

14. Measure absorbance at 570 nm and calculate percent activity.

2.5.4 Shipping Peptide Libraries

Collaborations with the Shaw Lab at USF and the Hoskin Lab at Dalhousie University have been established for ESKAPE pathogen screening and cancer screening, respectively. As such, peptide libraries must be shipped to each location. For best results, it is crucial that all analysis is done with the same peptide library on the same time frame to minimize the possibility of peptide degradation. As such, two peptide libraries from the same species should be fractionated on the same day and then combined to obtain 500 μL of each fraction. This volume is sufficient for either a full panel of ESKAPE pathogens or anticancer assays and half of the ESKAPE pathogens. If the full analysis is desired, 3 or 4 peptide libraries should be combined before analysis. Combined library should then be aliquoted into separate microtiter plates with the appropriate volume for analysis at each location (e.g. one plate for *E. coli* bioactivity and LC-MS analysis with 80 μL/well and one plate for ESKAPE pathogen screening with 420 μL/well). Plates being sent to collaborators should be tightly sealed using a well plate matt and stored at -20℃ until shipment. Plates should be shipped on a weekday (Monday – Thursday) via FedEx on dry ice with overnight delivery.

2.6 LC-MS/MS Analysis

LC-MS/MS analysis is used to obtain peptide abundance data across fractions 11 – 40 in each peptide library. Early and late-eluting fractions are excluded to avoid instrument contamination via highly abundant small molecules and high salt concentrations, respectively. For most peptide libraries, a dilution ratio of 5 μL of peptide fraction into 45 μL of 95% water/5% ACN/0.1% formic acid provides the appropriate loading level (low e7 TIC) for LC-

MS. However, it is recommended to pick a representative fraction based off 280 nm SCX chromatogram (usually between 18 and 25), start with a lower (5:100) dilution ratio, and increase concentration 2-fold until the optimal loading level is achieved. Once the desired dilution ratio is determined, prepare all of the remaining fractions at the same concentration (crucial for relative quantification across fractions) in a low-volume 96-well plate and cover with an adhesive plate seals.

Fractions are analyzed using a Waters nanoAcquity UPLC coupled to a Sciex TripleTOF 5600 mass spectrometer, following the parameters outlined in Tables 2-1 and 2-2. Five μL of each sample is injected onto a Waters Symmetry C18 trap column (5 μm particles, 180 μm ID, 20 mm length) before subsequent passing onto a Waters HSS T3 C18 column (1.8 μm particles, 75 μm ID, 250 mm length) analytical column. The LC employs a two-step linear gradient of increasing mobile phase B (step 1: 5-50%, step 2: 50-85%) over 60 minutes for peptide separation (Table 2-1). Electrospray ionization is used to generate gaseous ions. LC-MS/MS data is collected in positive ion, high sensitivity mode, using information dependent acquisition (IDA) of the first 20 ions for MS/MS. A BSA calibration run should be included after every 8 samples to ensure high mass accuracy.

All LC-MS data is processed using Progenesis QI for Proteomics software (Waters). Progenesis QI is a label-free proteomics software that enables peak alignment and quantification over multiple LC-MS runs. To align runs, allow the software to choose a reference run from a select subset of fractions (usually between 15 and 35) that elute over the SCX gradient and contain the region of activity. Apply automatic processing settings with a retention time filter using a minimum peak with of 0.1 minutes to align and peak pick ions across all runs. Peaks that are aligned proceed to be quantified using the survey scan data. Because of alignment of ions

across all fractions, we are able to track changes in abundance of each individual ion detected across SCX fractions and thus relate these changes to the changes in bioactivity. All data should be exported at the 'peptide ion' stage.

2.7 PepSAVI-MS Statistics

Exported peptide ion data is subject to PepSAVI-MS statistical analysis for determination of most-likely components contributing to the observed bioactivity. PepSAVI-MS statistics are implemented in R and have been developed into a publicly available CRAN package, PepSAVIms (https://cran.r-project.org/package=PepSAVIms). Development of the PepSAVIms statistical analysis package was performed in collaboration with the Liu Lab at UNC Chapel Hill. Graduate student David Pritchard wrote the code behind each function and helped to author the package vignettes, which serve as the basis of this section.

The PepSAVIms statistical analysis pipeline consists of three core functions: binMS, filterMS, and rankEN. The prototypical workflow for analysis in our laboratory is highlighted by the solid lines in Figure 2.2, using all three functions, in order, to extract a ranked compound list. However, the binMS and filterMS functions are optional workflow steps that may not be necessary for data generated from different mass spectrometers utilizing different peak picking software. Instead, this data can be imported and organized using the msDat feature before rankEN analysis. In all cases, bioactivity data is required at the rankEN step.

2.7.1 Binning of Mass Spectrometry Data: Using the *binMS* Function

Binning of mass spectrometry data is used to refine the data matrix generated from peak picking software. As PepSAVI-MS relies on the quantification of all species regardless of abundance or associated MS/MS information, it is ideal for all species to be picked. One way to

achieve this is to increase the sensitivity of the peak picking software; however, it has been observed that this often results in the same compound being 'picked' multiple times due to shifts in retention time or mass accuracy across fractions. As such, the binMS feature was designed to identify and group compounds designated as different using the peak picking software, but seem to be the same after accounting for mass accuracy and retention time figures of merit associated with the available LC-MS/MS system.

BinMS uses *m/z*, charge, and retention time parameters to determine compounds that should be binned together, according to the following notation:

binMS (mass_spec, mtoz, charge, mass, time_peak_reten, ms_inten, time_range, mass_range, charge_range, mtoz_diff, time_diff)

The mass spectrometry data matrix must be loaded as a .csv file into R, and is called into the binMS feature as 'mass_spec'. 'Mtoz', 'charge', 'mass', 'time_peak_reten', and 'ms_inten' values are parameters present in the mass spectrometry data matrix and are can be identified by the binMS function by column number or a unique column identifier. Retention time, mass, and charge filters can be employed using the 'time_range', 'mass_range', and 'charge_range' parameters to exclude compounds eluting during wash/equilibration or compounds too large or small to be considered bioactive peptides. The last 2 parameters of the binMS feature specify the requirements for two rows to be binned together. 'Mtoz_diff' specifies the largest allowed difference in two *m/z* ratios for the features to be considered for binning. Similarly, the 'time_diff' parameter specifies the largest allowed shift in retention time for two compounds to be considered the same.

Once all parameters are specified, the binMS feature begins by identifying the *m/z* ratio of the first compound and comparing it to every other feature in the spreadsheet. If both the 'mtoz_diff' and 'time_diff' criteria are satisfied, and the charge state of the two features is identical, the two rows will be binned together, averaging all parameters associated with compound identification and summing intensity values. This newly binned row is then compared to all other rows in the spreadsheet until no further binning can occur. Then, the next feature is selected for binning until all features have been considered. The resulting binned rows can be written into a new data matrix and exported as a .csv file. For our LC-MS platform, binning with the thresholds of 0.05 *m/z* and 5 minutes retention shift generally produces accurate binning. However, data should be manually interrogated to investigate the optimal binning thresholds for each data set.

2.7.2 Filtering of Mass Spectrometry Data: Using the *filterMS* Function

Binned mass spectrometry data is then filtered for compounds likely contributing to the observed bioactivity profile(s) using the 'filterMS' function, as follows:

filterMS (msObj, region, border, bord_ratio, min_inten, max_chg)

If binning was used, the output .csv should be the input mass spectrometry object, 'msObj'. However, mass spectrometry can enter directly into the filterMS function if desired using the MsDat function for formatting. 'Region' specifies the filtering region, as designated by the observed bioactivity profile. As mass spectrometry analysis is more sensitive than bioactivity measurements, filtering should allow for the active compound to be present in $\pm$ two fractions on either side of the bioactivity profile, designated as the extended bioactivity region. Typically, all other fractions are considered the 'border', however some could be excluded from this parameter

if desired. The ‘bord_ratio’ specifies the maximum allowed abundance of a compound outside of the bioactivity region, and is relative to the maximum observed abundance of that feature, only. To eliminate noise, peak picked data can also be filtered through requirement of a minimum intensity value or maximum charge value, as designated by ‘min_inten’ and ‘max_charge’.

FilterMS uses the input mass spectrometry data to filter out any compounds unlikely to be causing the observed bioactivity using the idea that if a compound is causing the activity, it should be maximally abundant in the fractions where bioactivity is observed (controlled by the ‘region’ and ‘border’ parameters). Following the same logic, that compound should not be highly abundant in any outside fractions where bioactivity is not observed (controlled by the ‘bord_ratio’ parameter). For the typical PepSAVI-MS dataset, the extended bioactivity region is most commonly between 5 and 7 fractions, and border ration of 0.05 (5%) is used. All compounds meeting the filterMS criteria are written into a new .csv file for entry into the rankEN function.

2.7.3 Penalized Linear Regression for Putative Bioactive Peptide Determination: Using the *rankEN* Function

Filtered compounds are subject to penalized linear regression to determine features with profiles most-closely matching the observed bioactivity profile using the rankEN function:

rankEN (msObj, bioact, region_ms, region_bio, lambda, pos_only, ncomp)

The output from the filterMS function serves as the ‘msObj’ input for the rankEN function. Bioactivity data across all fractions with corresponding mass spectrometry data is required to implement this function of the PepSAVIms package, and is called to the rankEN function as ‘bioact’. The ‘region_ms’ and ‘region_bio’ parameters specifications designate the mass spectrometry and bioactivity data to be compared, and can be called by column number or a

unique column identifier. Lambda specifies the penalty parameter to be used in the elastic net estimation, typically 0.01 for PepSAVI-MS data sets. The 'pos_only' feature allows only compounds with a positive correlation to be considered for modeling, and 'ncomp' specifies the maximum number of compounds the function should report.

The rankEN function employs penalized linear regression with an elastic net estimator to determine compounds most-likely contributing to the observed bioactivity. Penalized linear regression is a regression method that imparts a penalty each time a compound enters the model, with the ultimate goal of obtaining a subset of compounds that minimizes the loss function. The rankEN function works by selecting a choice of quadratic penalty parameter, and for fixed value of quadratic penalty parameter tracking the order in which the coefficients corresponding to candidate compounds first become nonzero along the elastic net path as the $\ell 1$ penalty parameter changes[4]. It is presumed that compounds with corresponding coefficients that become nonzero earlier in the path may be better candidates for having an effect on bioactivity levels then those with corresponding coefficients that become nonzero later in the path. A small, nonzero value is recommended for the quadratic penalty parameter in this instance. This value will approach the lasso penalty while bypassing the lasso limitation of selecting no more than one less than the number of bioactivity data points of nonzero coefficients, thus severely reducing the size of the list of candidate compounds[5]. In the end, the rankEN function returns identifying information for each compound in the order in which the corresponding regression coefficient for a given compound first becomes nonzero as part of the Elastic Net path.

2.7.4 General Usage

An example of the fully coded PepSAVIms statistical analysis is depicted in Figure 2.3. The PepSAVIms package, along with its supporting packages (devtools, lars, and elastic net), must

be loaded for use of core functions. Blue colored files serve as input files for each function, and orange files as the corresponding output. Optional functions are depicted in green. These functions are useful to ensure that data is being read in to R properly, and to display column headers useful for designating column numbers within each function. However, these functions are not required for package use. Each core function is equipped with a corresponding summary function that displays the number of compounds entering and exiting the compound after each round of analysis. It is highly advised to write the compounds remaining after each stage to a unique .csv file to ensure proper function and for use in troubleshooting.

2.8 Peptide Prioritization and Target Identification

Top 20 ranked compounds are considered the initial candidates for identification and purification, with priority starting with the highest ranking compounds or compounds appearing in ranked lists for multiple target species. As no peptide deconvolution is used, each peptide charge state or adduct is considered a unique feature and is allowed to enter the model individually. While this compound redundancy limits the number of unique compounds that enter the model, identification of a single peptide with multiple charge states/adducts ranking in the top 20 compounds serves as an additional level of confidence and can be used for further peptide prioritization.

Initial MS/MS analysis includes collision induced dissociation of the first 20 compounds meeting the fragmentation criteria. While this generates some level of sequence information for compounds on the top 20 list, targeted fragmentation is typically required after peptide prioritization. This can be achieved using parallel reaction monitoring (PRM) or an information dependent acquisition (IDA) experiment with an inclusion list using the fraction with the highest abundance of each compound. As many organisms have unsequenced or only partially

sequenced genomes, *de novo* sequencing is used for natural product peptide identification. *De novo* sequencing of each compound is performed using the *de novo* graphical user interface (GUI) in conjunction with manual sequencing[6]. If only partial sequences can be obtained, sequence tags can be searched against a related species to serve as a starting point for peptide assignment.

2.9 Peptide Isolation

Target peptides must be isolated or synthesized for activity validation and exploration of activity against a complete bioactivity panel. Peptide synthesis is the preferred route forward for linear, unmodified peptides with complete sequences. To enable production of peptides and small proteins, Fmoc-based solid-phase peptide synthesis (SPPS) will be performed in collaboration with the Womack Lab at High Point University using a flow chemistry-based platform built in-house[7-8]. Using this proven instrument, peptide fragments are rapidly generated and further purified with modern chromatography methods. If sequence determination is difficult or ambiguous, or if it is not feasible to synthesize the identified peptide, isolation from source material will be conducted using the applicable chromatographic chemistry. While this approach is more labor intensive and will require chromatographic optimization, it allows for the isolation of complex/difficult-to-sequence peptides.

2.10 Minimum Inhibitory Concentration (MIC) Determination

Purified peptide(s) will be subject to an array of relevant bioactivity screens determined by the initially observed activity. Minimum inhibitory concentrations will be determined using a dilution series of isolated peptide against each target pathogen. Minimum inhibitory concentrations can be determined in either the microbroth dilution or agar diffusion-based assays.

2.11 Tables

Table 2-1 Liquid chromatography conditions.

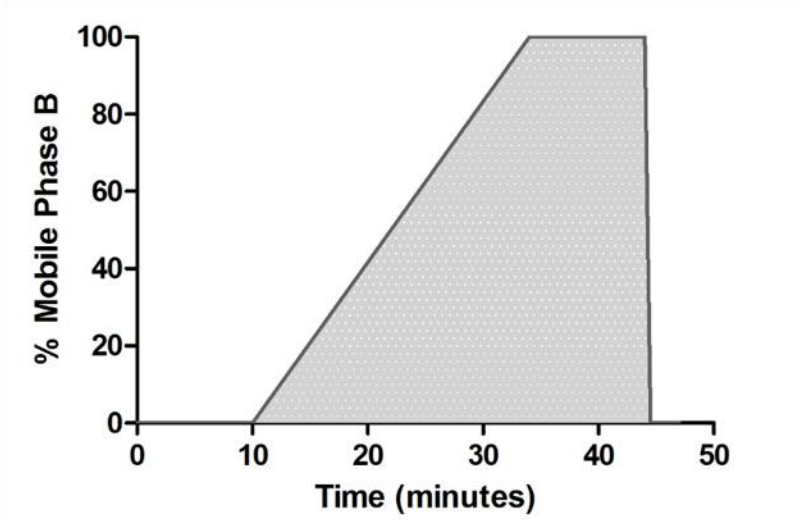
LC conditions:		
LC system	Waters nanoAcquity	
Mobile phase A (MPA)	Water with 0.1% formic acid	
Mobile phase B (MPB)	Acetonitrile with 0.1% formic acid	
Trap column	Waters Symmetry C18 trap column (5 μm particles, 180 μm ID, 20 mm length)	
Trapping conditions	5 uL/minute, 100% MPA for 3 minutes	
Analytical column	HSS T3 C18 column (1.8 μm particles, 75 μm ID, 250 mm length)	
Flow rate	0.3 μL/min	
Run time	60 minutes	
Gradient:		
Time (minutes)	% MPA	% MPB
0	95	5
1	95	5
31	50	50
33	15	85
36	15	85
37	95	5
60	95	5

Table 2-2 Mass spectrometry acquisition parameters.

Mass spectrometry acquisition parameters:	
MS system	AB Sciex TripleTOF 5600
Ionization	Electrospray
Operation	Positive-ion, high sensitivity mode
Survey spectrum (MS^1) mass range	350 – 1600 Da
Survey spectrum acquisition time	250 ms
MS^2 mode	Information-dependent acquisition (IDA)
MS^2 selection criteria	First 20 features >150 counts/second with charge state +2 to +5
Collision energy	Rolling ($\pm$ 5%)
MS^2 exclusion window	8 seconds
Total cycle time	2 seconds
Run time	60 minutes

2.12 Figures

a.



b.

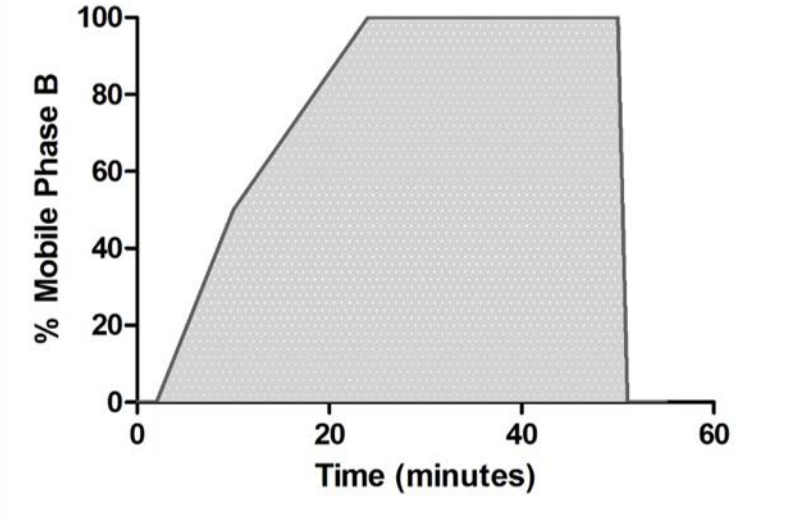


Figure 2.1 Optimized SCX gradients for (a) plant extracts and (b) microbial secretome extracts.

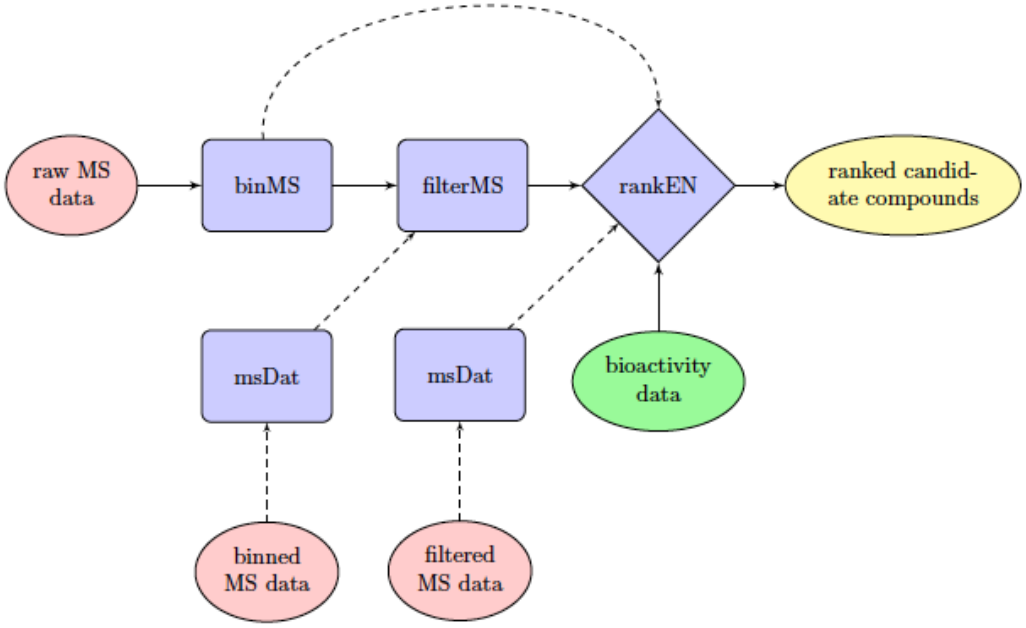


Figure 2.2 PepSAVIms statistical analysis workflow, adapted from the PepSAVIms CRAN package vignette. Solid lines represent the prototypical workflow for the Hicks Laboratory, and dashed lines represent optional data entry and analysis routes.

```
setwd("/Users/Christine/Documents/Hicks Lab/Statistical modeling/20180227_EPu_SCX2 modeling")

library("devtools", "lars", "elasticnet", "PepSAVIms")

MSdata<-read.csv("20180227_CLK_BAP_EPu_4x_SCX2_fractions_peptide ion data_clean.csv")

names(MSdata)

MSdata[1:5, 1:5]

binMS_out<-binMS(MSdata, 1, 4, 3, 2,5:34, c(14,45), c(1000,10000), c(2L,10L), 0.05, 5)

summary(binMS_out)

bin_matrix<-extractMS(binMS_out)

write.csv(bin_matrix,"20180227_EPu_SCX2_binned.csv")

filterMS_out<-filterMS(binMS_out, paste0("SCX2_",25:31), border = "all", bord_ratio = 0.1, min_inten = 1000,
max_chg = 10L)

summary(filterMS_out)

filter_matrix<-extractMS(filterMS_out)

write.csv(filter_matrix,"20180227_EPu_SCX2_filtered.csv")

MSdata1<-read.csv("20180227_EPu_SCX2_filtered.csv")

names(MSdata1)

rawMsDat<-msDat(MSdata1, 2, 3, ms_inten=4:33)

bioactivity<-read.csv("20171115_EPu 4x bioassay_data only.csv")

bioactivity

dim(bioactivity)

E.coli      <- bioactivity[1L, ]

save(rawMsDat,

    E.coli,

    file="20180227_EPu_SCX2_combined MS and bioactivity.RData")

load("20180227_EPu_SCX2_combined MS and bioactivity.RData")

orderedCompds_E.coli<-rankEN(rawMsDat,E.coli, paste0("SCX2_",26:30), paste0("SCX2_",26:30),
0.001,pos_only = TRUE, ncomp = NULL)

summary(orderedCompds_E.coli)

write.csv(extract_ranked(orderedCompds_E.coli, TRUE),"20180227_EPu_SCX2_ranked.csv")
```

Figure 2.3 Example R code for PepSAVI-MS statistical analysis, considering typical workflow processing and a bioactivity region of fractions 26-30.

REFERENCES

1. Hammami, R.; Ben Hamida, J.; Vergoten, G.; Fliss, I., PhytAMP: a database dedicated to antimicrobial plant peptides. *Nucleic Acids Res.* **2009,** *37* (Database issue), D963-968.

2. Sarker, S. D.; Nahar, L.; Kumarasamy, Y., Microtitre plate-based antibacterial assay incorporating resazurin as an indicator of cell growth, and its application in the in vitro antibacterial screening of phytochemicals. *Methods* **2007,** *42* (4), 321-324.

3. Leslie, J. F.; Summerell, B. A., *The Fusarium laboratory manual*. Blackwell Publishing: Ames, Iowa, **2006**.

4. Zou, H.; Hastie, T., Regularization and variable selection via the elastic net. *J. Roy. Stat. Soc. Ser. B. (Stat. Method.)* **2005,** *67* (2), 301-320.

5. Tibshirani, R., Regression Shrinkage and Selection via the Lasso. *J R Stat Soc Series B Stat Methodol.* **1996,** *58* (1), 267-288.

6. Muth, T.; Weilnbock, L.; Rapp, E.; Huber, C. G.; Martens, L.; Vaudel, M.; Barsnes, H., DeNovoGUI: an open source graphical user interface for de novo sequencing of tandem mass spectra. *J. Proteome Res.* **2014,** *13* (2), 1143-1146.

7. Simon, M. D.; Heider, P. L.; Adamo, A.; Vinogradov, A. A.; Mong, S. K.; Li, X.; Berger, T.; Policarpo, R. L.; Zhang, C.; Zou, Y.; Liao, X.; Spokoyny, A. M.; Jensen, K. F.; Pentelute, B. L., Rapid flow-based peptide synthesis. *Chembiochem : a European journal of chemical biology* **2014,** *15* (5), 713-720.

8. Mijalis, A. J.; Thomas, D. A., 3rd; Simon, M. D.; Adamo, A.; Beaumont, R.; Jensen, K. F.; Pentelute, B. L., A fully automated flow-based approach for accelerated peptide synthesis. *Nature chemical biology* **2017,** *13* (5), 464-466.

CHAPTER 3: Development and Validation of the PepSAVI-MS Pipeline for Natural Product Bioactive Peptide Discovery

3.1 Introduction

Difficult-to-treat nosocomial and community acquired fungal and extensively drug resistant bacterial infections are increasingly commonplace, new viral diseases are emerging and spreading rapidly, and cancer remains a leading cause of death worldwide. In the US, there are almost two million hospital-acquired bacterial infections each year, resulting in ~100,000 deaths[1]. The emergence of ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* species) pathogens has led to a significant increase in multidrug resistant (MDR) infections in the clinic with associated increases in morbidity and mortality[2]. Additionally, the recent emergence, resurgence, and spread of viruses, including Zika, SARS, West Nile, Ebola, and MERS[3], for which there are limited or no direct-acting antivirals, highlight the susceptibility of the human population to future potentially untreatable pandemics. Furthermore, despite continued progress in anticancer therapeutics, limitations to current lead compounds of nonspecific toxicity, poor drug penetration, and multidrug resistance have emphasized the need for discovery of anticancer therapeutics with novel mechanisms of action[4]. With many bacteria now unresponsive to multiple classes of antimicrobial compounds and cancer the leading cause of death worldwide,

there is an undeniable and desperate need to identify novel pharmacological chemistries and accelerate their development through new methods and innovative technologies.

Natural products have long been sources of virtually all traditional medicinal preparations and have been the single most successful source of lead compounds for drug discovery[5]. Specifically, plants have played a significant role in the treatment of human ailments since prehistoric times. The teas and tinctures of times past are one source to drug discovery, allowing the ethnobotanically-guided isolation and characterization of pharmacologically active compounds for the treatment of bacterial and fungal infections, cancers, and other ailments. Despite historical relevance and past success, challenges associated with natural product discovery have slowed progress in drug discovery efforts. Additionally, natural product discovery efforts have largely focused on small molecule constituents. However, recent discoveries have revealed ribosomally-synthesized, post-translationally-modified peptide natural products (RiPPs) with substantial structural diversity and bioactivity potential, including novel mechanisms of action[6-7]. While traditionally studied for antifungal and antibacterial properties, recent studies have piqued interest in these peptides as potential anticancer therapeutics[8]. Peptides offer several advantages over other small and large molecule drug candidates – including greater efficacy, selectivity and specificity relative to small organic molecules, better tissue penetration, and reduced immunogenicity and manufacturing costs relative to proteins/antibodies. Advances in peptide modification, formulation, and delivery methods can address known limitations of peptidic drug candidates[9], including modification of peptide length/content to increase selectivity[10], stapling and/or peptidomimetic conversion techniques to improve pharmacokinetic properties[11], as well as encapsulation methods to protect from proteolytic degradation[10]. However, RiPPs fall outside the scope of standard therapeutic

screening approaches and elude detection via standard -omic workflows due to their large size, structural diversity and high level of post-translational modification; therefore, systematic approaches for discovery and characterization are in nascent development[12]. Developing natural product screens for antimicrobially active RiPPs has the potential for discovery of new bioactive compounds with novel mechanisms of action able to address the growing problem of antimicrobial resistance.

Current methods for RiPP discovery often rely on bioassay guided fractionation[13] or genomic mapping[12, 14-15] (e.g. Pep2Path, RiPPquest) to facilitate downstream analysis. Relying on iterative rounds of chromatographic separations, bioassay guided fractionation is extremely time consuming and provides no structural information until late in the discovery process. Additionally, this approach often leads to replication of previously known compounds as a bias towards highly abundant and highly active compounds is evident. While alternative genomic mapping approaches can provide structural information from the beginning, prior genomic sequencing and knowledge of peptide biosynthetic pathways is required. Furthermore, this approach is unable to provide direct bioactivity information on the target peptide thus necessitating downstream isolation and activity screening to determine function and biological activity. Recent approaches to address some of these limitations have been proposed[16] (e.g. Compound Activity Mapping); however, these platforms are specific to small molecule identification and do not translate efficiently towards advancing untargeted proteomics approaches.

To this end, we have developed the PepSAVI-MS (Statistically-guided bioactive peptides prioritized via mass spectrometry) pipeline for the screening and identification of cationic bioactive peptides from natural product sources. For streamlined identification of bioactive

peptides, our platform employs only one round of crude fractionation and relies on the power of mass spectrometry and statistics to assign bioactivity to individual components. Furthermore, we implement physiologically relevant whole cell bioassays to obtain activity information early in the discovery process, thus focusing on only active bioactive peptide targets. Our PepSAVI-MS pipeline is adaptable to any natural product source of peptides and can test against multiple physiological targets of various cell lines or organisms, including bacteria, fungi, viruses, protozoans, and cancer cells. Herein, we demonstrate the PepSAVI-MS pipeline using plant material and focusing on peptides with antimicrobial and/or anticancer properties. Proof-of-principle studies for a known AMP from the plant *Viola odorata* are presented to validate this workflow. Additionally, we use this pipeline to explore novel antibacterial and anticancer activities of the *Viola odorata* cyclotide, cycloviolacin O2.

3.2 Materials and Methods

3.2.1 Source Materials and Growth Conditions

Viola odorata seedlings purchased from Strictly Medicinal Seeds (formerly Horizon Herbs; Williams, OR) were grown to mature rosette stage (~6 weeks) in standard greenhouse conditions. Seedlings were planted in nutrient-rich soil under controlled temperature (63.5 - 68.5°F) and light cycle (14-hr light) conditions. Aerial tissue was harvested with immediate flash freezing under liquid nitrogen, and stored at -80°C until subsequent extraction.

3.2.2 Creation of Peptide Libraries

Extraction. Frozen tissue (100 g) was ground under liquid nitrogen using a mortar and pestle and aqueous extracted in 300 mL of 10% acetic acid with protease inhibitors (Roche, 1 tablet/50 mL) and stirring for 4 hours at 4°C. Insoluble material was pelleted by centrifugation at 13,000 rcf for 45 minutes, and 0.45 μm stericup filtration (Millipore) was used to remove remaining

particulates. Protein concentration filters (MWCO 30kDa; Millipore) were used to remove high molecular weight species, and subsequent dialysis (0.1 – 1 kDa cutoff; SpectrumLabs) into 5 mM ammonium formate pH 2.7 was performed to eliminate small molecule contaminants. The sample was concentrated via vacuum centrifugation to 4 mL to generate the final crude extract.

HPLC fractionation. Extracts were subjected to a 47-minute SCX method using a PolySulfoethylA column (100 x 4.6 mm, 3 μm particles, PolyLC). A salt gradient (Figure 3.1) was employed using a linear ramp from 5 mM ammonium formate, 20% acetonitrile, pH 2.7 to 500 mM ammonium formate, 20% acetonitrile, pH 3.0. Fractions were collected across the 47-minute run in one-minute intervals and desalted with three washes of 1.3 mL water using a vacuum concentrator. Peptide libraries were stored in water (400 μL/fraction) at 4°C until bioactivity assay.

3.2.3 Bioactivity Screening

Bacterial. Escherichia coli (25922) was obtained through ATCC and the ESKAPE pathogen strains are clinical isolates that belong to a collection acquired by the Shaw Lab [*Enterococcus faecium* (1449), *Staphylococcus aureus* (635), *Klebsiella pneumoniae* (1433), *Acinetobacter baumannii* (1403), *Pseudomonas aeruginosa* (1423), *Enterobacter cloacae* (1454)][17]. Bacterial cultures were synchronized to mid-log phase by inoculating 1 mL of overnight culture into 100 mL of TSB and incubated for 3 hours, shaking at 37°C. All assays were performed in 96-well microtiter plates by adding 10 μL 2x MH broth, 20 μL of 1x MH broth, 10 μL peptide fraction, and 10 μL 0.5 OD_{600} bacteria culture. The addition of 2x broth ensured sufficient nutrients were available since the peptide fractions, in water, accounted for 1/5 of the final volume. TSB was used to test *S. aureus* due to its decreased growth rate in MHB. Ampicillin (100 μg/mL) or Tetracycline (100 μg/mL) was added respectively to *E. coli* or ESKAPE pathogen plates in place

of the plant fraction to serve as the positive control, and water as the negative control. The prepared plates were incubated at 37°C, 275 rpm shaking for 1.5 - 4 hours depending on the growth characteristic of the organism before 5 µL of the cell viability indicator dye, resazurin (1.19 mM), was added to each well. After one additional hour of shaking and incubation, a fluorescence read of 544 nm (ex) and 590 nm (em) was collected to measure relative fluorescence units for each sample. Percent activity of each well was calculated using the formula: $\% \, Activity = \left(1 - \left(\frac{RFU \, of \, fraction - RFU \, of \, positive \, control}{RFU \, of \, negative \, control - RFU \, of \, positive \, control}\right)\right) * 100\%$.

Cancer cell lines: Cell viability assay. MDA-MB-231 breast cancer cells, OVCAR-3 ovarian cancer cells, PC3 prostate cancer cells, or human dermal fibroblasts in the appropriate growth medium were plated into 96-well flat-bottom microtiter plates at 2×10^4 cells/well and cultured for 24 hours in a 37°C humidified incubator in order to form adherent monolayers. Medium was then removed and replaced with fresh serum-free medium with or without individual plant fractions to be assayed for cytotoxic activity. Cells were then cultured for an additional 24 hr. Two hours before the end of culture, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) solution was added to each well at a final concentration of 0.5 mg/mL to measure mitochondrial succinate dehydrogenase activity. Culture supernatants were discarded and formazan crystals were solubilized in 0.1 mL of dimethyl sulfoxide. Absorbance was measured at 570 nm on an ASYS Expert 96 plate-reader (Montreal Biotech Inc., Kirkland, QC). Absorbance in each treatment group was compared to the control to determine percent reduction in viability, as reflected by the change in mitochondrial succinate dehydrogenase activity.

3.2.4 LC-MS/MS

Peptide libraries were analyzed via a nano-LC-ESI-MS/MS platform: nanoAcquity (Waters, Milford, MA) coupled to a TripleTOF5600 (AB Sciex, Framingham, MA). Peptide fractions were diluted to the appropriate loading level, acidified with formic acid and transferred to low-volume 96-well plates covered with adhesive plate seals. Each sample was injected onto a trap column (NanoAcquity UPLC 2G-W/M Trap 5 µm Symmetry C18, 180 µm x 20 mm: Waters) before subsequent passing onto the analytical C18 column (10k PSI, 100 Å, 1.8µm, 75µm x 250 mm: Waters). Peptide separation was carried out at a flow rate of 0.3 µL/minute using a linear ramp of 5 – 50 % B (mobile phase A, 1% formic acid; mobile phase B, 1% formic acid in acetonitrile) over 30 minutes. The MS was operated in positive ion, high sensitivity mode with the MS survey spectrum using a mass range of 350-1600 Da in 250 ms and information dependent acquisition (IDA) of MS/MS data. The first 20 features above 150 counts threshold and having a charge state of +2 to +5 were fragmented using rolling collision energy $\pm$5%. Auto calibration was performed every eight samples (8 hours). The data have been deposited to the ProteomeXchange Consortium via the PRIDE[18] partner repository with the dataset identifier PXD004780. De-isotoped peak lists for each fraction were generated using Progenesis QI for Proteomics software (Nonlinear Dynamics, v.2.0). To align runs, a reference run was chosen from a subset of bioactive fractions (15 - 30 for *V. odorata*). Automatic processing settings with a retention time filter of 14 - 45 minutes were used to align and peak pick ions across all runs. Identified features were quantified using AUC integration of survey scan data based on the summed intensity of each de-isotoped feature. Data was exported as "peptide ion data" with the default parameters from Progenesis at the "Identify Peptides" stage in the software. Our analysis yielded 6,258 unique MS features for *V. odorata*.

3.2.5 Data Reduction and Statistical Modeling

Areas of interest in the bioactivity profile were selected for subsequent data reduction and modeling. The bioactivity region for each *V. odorata* data set was defined for each pathogen based on the observed bioactivity profile as follows: ovarian cancer (18-22), breast cancer (18-22), prostate cancer (18-22), *A. baumannii* (17-21), *E. faecium* (17-21), *P. aeruginosa* (18-21), *E. coli* (18-25), *F. graminearum* (19-24). Using the PepSAVI-MS software, background ions were eliminated through retention time, mass, and charge state filters to reduce the data to potential compounds of interest. Peak-picked data believed to belong to the same MS feature were binned together if their *m/z* values were within a 0.05 Da window and their charge states were identical. Next, workflow-informed filtering criteria were applied using the following rules: 1) *m/z* intensity maximum must fall inside the range of the bioactivity "area of interest", 2) the *m/z* intensity of species meeting the first criteria must be <1% of its respective maximum peak intensity in the areas bordering said "area of interest", 3) there must be non-zero abundance in the fraction following the maximum intensity fraction, 4) the maximum intensity must be > 1,000 in active window, 5) all charge states > +10 are excluded. All *m/z* species meeting these filtering criteria were modeled using the elastic net estimator with a quadratic penalty parameter specification of 0.001 to determine each species' contribution to the observed overall bioactivity profile. The resulting list contains candidate compounds ranked in order of when they entered the model, such that the highest ranked compounds have the greatest likelihood to be contributing to the bioactivity.

3.2.6 Isolation of cyO2

Dried aerial plant material from *Viola odorata* (Alfred Galke GmbH, Bad Grund, DE) was used to isolate cyO2. Cyclotide isolation was carried out according to Herrmann *et al*. with

minor modifications[19]. Briefly, ground plant material was subjected to extraction with 60% methanol in water, followed by filtration and liquid–liquid extraction with dichloromethane. The polar phase was diluted and acidified before subjected to a strong cation exchange chromatography to capture cyclotides of net positive charge. Individual cyclotides, i.e. cyO2, were then isolated using RP-HPLC. Final cyO2 was of >95% purity assessed by RP-HPLC (214 nm).

3.2.7 Validation of cyO2 Activity

Minimum inhibitory concentration determination of cyO2 against *E. coli* and *A. baumannii* was performed in Mueller-Hinton media using a broth microdilution method. Bacteria were grown overnight and added to the assay at 10^5 CFU/mL and turbidity/clearing was observed after 20 hours of incubation. Assay was performed in triplicate starting with 10 or 25 μM cyO2 for *E. coli* or *A. baumannii*, respectively, and using 2x serial dilution. Mean MICs of 5 and 15 μM were obtained for each species, respectively. The mean IC_{50} value of 6.6 μg/mL for cyO2 against the PC3 prostate cancer cell line was determined using duplicate measurements of the MTT cell viability assay as described above.

3.3 Results and Discussion

3.3.1 Overview

Relying on a multi-pronged approach, the PepSAVI-MS pipeline utilizes selective extraction and fractionation of peptide source material, bioactivity screening, and statistics-guided mass spectrometry-based peptidomics for the targeted identification and characterization of only putative bioactive compounds (Figure 3.2). We demonstrate use of the PepSAVI-MS pipeline

focusing bioactive peptides from the plant kingdom and their activity against a panel of microbial pathogens and cancer cell lines.

3.3.2 Plant Selection, Cultivation, and Harvesting

Plant species are selected based on several criteria including known bioactivity of the plant, traditional use, and availability. While many plants are commonly used for their perceived health-related benefits, the bioactive peptides from these species have not been comprehensively evaluated. Seeds/seedlings of selected species are obtained through commercial sources and grown in a laboratory greenhouse. Tissue is harvested with immediate flash freezing in liquid nitrogen for future extraction.

3.3.3 Peptide Library Creation

Preparation of crude extracts from plant material includes aqueous extraction and size exclusion to selectively target AMP-like molecules, i.e. water-soluble compounds that are smaller than 10 kDa[20-21]. Extracts are crudely fractionated using strong cation exchange (SCX) chromatography such that individual peptides elute over multiple sequential fractions. The distribution of bioactive peptide abundance across sequential fractions is reflected in the distribution of activity seen in bioactivity data for the same fractions. As peptides are eluted using volatile salts during SCX, peptide libraries are simply desalted via vacuum concentration for compatibility with downstream bioassays (Figure 3.2a). Peptide libraries can be created from constitutive expression, as well as from abiotic and biotic stresses known to induce production of defense compounds[22-25].

3.3.4 Bioactivity Screening

Peptide libraries are assayed for growth inhibition against pathogens of interest using physiologically relevant whole-cell bioactivity screens (Figure 3.2b). Library fractions are

incubated with a microbial or cancer cell culture and the presence of bioactive peptides in a given fraction will result in inhibition of culture growth during the incubation period. For bacterial assays, the remaining viable cells are quantified indirectly by spectrophotometric measurement of the irreversible intracellular reduction of resazurin[26]. For anticancer bioactivity, cytotoxicity assays are performed using MTT-based assays to measure mitochondrial succinate dehydrogenase activity with absorbance measurements at 570 nm. Values for each fraction are compared to positive and negative controls containing a known therapeutic or water, respectively, to determine a percent activity of each library fraction, where a small value of remaining viable cells indicates high activity. Species with bioactivity profiles of interest are prioritized for LC-MS/MS and downstream statistical analysis.

3.3.5 MS Profiling, Data Reduction, and Statistical Modeling to Identify Putative Bioactive Peptides

Peptide libraries are analyzed via nanoLC-ESI-MS/MS to obtain accurate intact mass and relative intensity information for peptide constituents contained within each library fraction (Figure 3.2c). The resulting data array is processed using the developed PepSAVI-MS software package. Using this function, unwanted compounds can be eliminated using mass, charge, and retention time filters. Remaining compounds are binned to condense all of the hits for a given *m/z* ratio into a single feature. Binned datasets are then filtered using the following workflow-informed criteria to narrow the library to those peptides most likely contributing to the bioactivity profile: 1) *m/z* intensity maximum must fall inside the range of the bioactivity "area of interest", 2) the *m/z* intensity of species meeting the first criteria must be <1% of its respective maximum peak intensity in the areas bordering said "area of interest", 3) there must be non-zero abundance in fraction following the maximum intensity fraction, 4) the maximum abundance must be > 1,000 in active window, 5) all charge states > +10 are excluded. This filtering results

in a reduction of each data set to those peptides with their highest abundance in the bioactive region. Subsequently, sparse penalized linear regression modeling (elastic net[27-28], including lasso[29]) is applied to correlate the relative abundance of post-filtered peptides and percent activity across library fractions in order to identify the specific peptide(s) causing bioactivity (Figure 3.2d). Because multiple peptides could jointly contribute to the bioactivity across a group of active fractions, a simple marginal correlation analysis strategy may produce inaccurate or spurious results for any individually contributing peptide. However, sparse penalized linear regression is able to account for this co-contribution to activity and results in more accurate identification of putative bioactive peptide species under the theoretical model. Elastic net regularization is an appropriate penalization specification as it elicits a sparse set of peptides contributing to bioactivity levels, helping to identify peptides most likely responsible for the corresponding bioactivities without strict limitation to a set number of peptides being selected. Identification of such a set facilitates prioritization of peptides for subsequent characterization and validation (Figure 3.2e). The elastic net path for fixed choice of quadratic penalty parameter can be obtained through a transformation of the predictor data and a subsequent application of the LARS algorithm[30]. Thus, in order to obtain a list of candidate compounds most likely to be contributing to the bioactivity, the order in which compounds' coefficients first obtain a nonzero value is tracked along the course of the algorithm, with coefficients that become nonzero earlier presumed to be better candidates for further investigation.

3.3.6 Peptide Characterization

Prioritized peptides are characterized via appropriate strategies depending on the target(s) of interest. While a top-down approach[31] is used for detection of peptides of interest and may be sufficient for characterization, a hybrid approach of chemical or enzymatic reaction followed by

LC-MS/MS analysis may be necessary to facilitate sequencing and identification of post-translational modifications (PTMs). The accurate intact mass and MS/MS spectrum of the peptide species is obtained from the initial LC-MS/MS data collection. If necessary, further characterization is carried out utilizing the library fraction containing the most abundant amount of the peptide species of interest. Reduction and alkylation of cysteine residues is used to determine the number of disulfide bonds. Proteolysis is used to assess disulfide bond connectivity and possible cyclizations (e.g. cyclotides containing a single glutamic acid residue can be linearized through cleavage with endoproteinase GluC[32]). Paired Lys-C and Lys-N digestions can facilitate determination of termini by allowing identification with high confidence of b- and y- ions from the isobaric precursor ions[33]. Complementary tandem MS fragmentation via CID and ECD/ETD is used in combination to maximize peptide primary sequence determination and post-translational modification localization[31, 34-35]. *De novo* peptide sequencing will be facilitated using PEAKS software (Bioinformatics Solutions Inc.)[36], supplemented with composition-based sequencing[37,38] and homology-based database searches[39] to address plant species with limited sequencing information available. Structural mapping to delineate disulfide bond connectivity is carried out in combination with multi-enzyme digestions[35], as appropriate, to dissect these key features that often confer bioactivity.

3.3.7 Isolation and In Vitro/In Vivo Validation of Prioritized Bioactive Peptides

Bioactive peptides that have been characterized are validated and their requisite bioactivity characterized. *In vivo* confirmation/validation of a bioactive peptide requires isolation to a high degree of purity to allow testing without interference from contaminating co-eluting compounds. Split-flow UPLC-MS or *de novo* synthesis of peptides will permit necessary biological and

pharmacological testing to determine if a peptide is a satisfactory lead compound that should be further evaluated.

3.3.8 Platform Validation

To validate the PepSAVI-MS pipeline, we demonstrate successful detection and identification of a known AMP from the botanical species *Viola odorata. Viola odorata*, commonly known as sweet violet, contains many cyclotides[40] - including cycloviolacin O2 (cyO2). CyO2 is a small, cysteine rich cyclotide comprised of 30 amino acids ($MW_{monoisotopic}$: 3138.37 Da), which has been shown to have diverse activity against many Gram-negative bacteria (*E. coli, K. pneumoniae*, and *P. aeruginosa*)[41], as well as several cancer cell lines[42-43]. Following the PepSAVI-MS pipeline as described above, *V. odorata* plants were grown, harvested, extracted, and crudely fractionated for creation of its peptide library. Using the known activities of cyO2 as a guide for assay selection, antibacterial bioassays were performed against *E. coli* (Figure 3.3a) and a panel of clinical strains representing the ESKAPE pathogens[17] (Figure 3.4). Each bioassay yielded robust bioactivity profiles unique to each microbe tested. Fractions containing the mass corresponding to cyO2 demonstrated growth inhibition for *A. baumannii*, *P. aeruginosa* and *E. coli*, but not for *E. faecium*, *S. aureus*, *K. pneumoniae*, or *E. cloacae*. As demonstrated previously[44], the crude nature of SCX separation often results in group isolation of cyclotides (Figure 3.5). Hence, these fractions also contained additional cyclotides that have the potential to be contributing to the aforementioned activity. The majority of the remaining fractions did not show growth inhibition, indicating the source of activity was due to constituents within the fractions rather than the extract itself. To further demonstrate the broad applicability of the platform, *V. odorata* fractions were screened for activity against human breast (MDA-MB-231), prostate (PC3), and ovarian (OVCAR-3) carcinoma cell lines and yielded robust activity

profiles with increased cytotoxicity for these neoplastic cells in comparison to a non-cancerous human dermal fibroblast cell line (Figure 3.6). *Viola odorata* peptide library was also screened for growth inhibition against the filamentous fungus, *Fusarium graminearum* (PH-1), which similarly yielded a robust activity profile across the cyclotide-containing fractions (Figure 3.7). To determine which of the detected cyclotides could be contributing to the aforementioned bioactivity, *V. odorata* peptide library was subjected to LC-MS/MS and statistical analysis with the developed PepSAVI-MS software package. Retention time alignment and peak-picking of detected ions identified 6,258 unique features across *V. odorata* fractions 11 – 43. To validate our automated workflow for data processing, extracted ion chromatograms of cyO2 peptide were plotted for manual versus automated data extraction (Figure 3.3b). Extracted ion chromatograms show that cyO2 was detected at varying abundances across fractions 18 – 22 and validate the use of automation for data array generation. Following the known mass and charge properties of bioactive peptides, exported peptide ion data was filtered to include masses between 2 and 15 kDa having charge states between +2 and +9 and eluting between 14 and 45 minutes. Remaining data was then binning using a 0.05 Da mass window with identical charge state requirements. Using the workflow informed criteria (described above) for data filtering and reduction with the bioactivity region specified as fractions 17 – 25, the number of possible candidates contributing to bioactivity was reduced to 225. This resulting list of candidate peptides was then subjected to statistical modeling of all bioactivity data sets, using the sparse penalized regression method, elastic net[28]. CyO2 from *V. odorata* was within the top 20 candidates for 2 of the 7 active species when modeled against the averaged bioactivity data. Furthermore, modeling individual bioactivity replicates with the requirement that any true component must be pulled out in at least two of the three replicates can further reduce the number of considered compounds for each data

set. CyO2 was characterized via a multi-step mass spectrometry-based approach (Figure 3.8) and these results are in agreement with previously published findings, thus confirming the identity of cyO2[40, 45]. MS^2 sequence coverage across the peptide allowed for distinction from another cyclotide, cyO9, present in sweet violet with the same intact mass as cyO2 and differing only by three residues. Only one other cyclotide corresponding to the intact mass of cyO17 appeared in the top 20 candidates after data reduction and statistical modeling of the *E. coli* bioactivity data set. No other known cyclotides appeared in the top 20 candidates for any of the other bioactivity data sets. While this does not eliminate the possibility that other cyclotides could be contributing to the activity, our experiments suggest that cyO2 was the main contributor. To allow for greater access to the scientific community, the statistical analysis methods established for this platform have been developed into an R package, PepSAVIms, that is publically available through CRAN (https://cran.r-project.org/package=PepSAVIms). Included with this package are two vignettes providing an in-depth description of the functionality of the PepSAVI-MS pipeline package, as well as the specific bioinformatics implementation performed in this manuscript.

To confirm that these bioassays are capable of detecting AMP activity at concentrations relevant for the observed activity, the minimum inhibitory concentration (MIC) of isolated cyO2 against *E. coli* was determined. Using a broth microdilution assay in 96-well plate format, the MIC was determined to be 5 μM, which lies in between the previously reported MIC of 2.2 μM[41] and MIC_{50} of 6.8 μM[13] for the same peptide on *E. coli* using different protocols. The concentration of cyO2 in the most abundant fraction (21) that killed or inhibited these strains was determined to be ~300 μM via area-under-the-curve integration of RPLC-MS peaks, which is in excess of the MIC needed for growth inhibition.

3.3.9 Novel Findings

In addition to platform validation, unknown activities of cyO2 from *V. odorata* have been revealed in this study. Previously, it was demonstrated that cyO2 was active against a subset of the ESKAPE pathogens, including *K. pneumoniae*, and *P. aeruginosa*, and inactive against the tested Gram-positive bacteria, including *S. aureus* and *E. faecium*[41]. Results from our study are in agreement with the activity of cyO2 in the cyclotide fractions against *E. coli* and *P. aeruginosa*, and inactivity against *E. faecium* and *S. aureus*. However, we could not detect activity in these fractions against *K. pneumoniae*, which is likely due to high strain variability among this bacterial species[46-47]. Additionally, our assay detects strong activity of the cyclotide fractions against *A. baumannii*, of which there are no known reported studies of activity (Figure 3.4). The MIC for the ESKAPE pathogen representing novel activity, *A. baumannii*, was determined to be 15 μM using isolated cyO2 and thus supports that cyO2 is the main contributor to the activity seen in these fractions. Throughout this study we have noticed variation in the sensitivity of peptidyl activity due to differences in media conditions and salt concentrations (Figure 3.9). We suspect that small differences even within different brands of the same medium could affect activity[48-49] necessitating validation of new lots of media through testing of established fraction libraries to confirm consistency with previously collected bioactivity data. While our study shows some examples of changes in activity due to strain variability, the majority of findings remain consistent across different strains of the same species.

Screening of *V. odorata* fractions against a panel of human cancer cell lines similarly demonstrates proof-of-principle, as well as an additional novel finding. In previous studies, cyO2 has been shown to have anticancer activity against multiple breast and ovarian cancer cell lines[42-43]. Although our screen used different cell lines, strong activity was observed against the MDA-

MB-231 breast cancer cell line, as well as the OVCAR-3 ovarian cancer cell line across fractions in which cyO2 was eluted. Additionally, this is the first demonstration of the ability of purified cyO2 to kill PC3 prostate cancer cells, with a determined IC_{50} of 6.6 μg/mL (2.1 μM). Furthermore, we demonstrate that the cyclotide fractions can kill the filamentous fungus *F. graminearium* and suspect the presence of an additional peptide(s) from *V. odorata* possessing antifungal capabilities (fractions 28 – 36). These studies not only highlight the wide applicability of this platform but also the capabilities of antimicrobial peptides as broad-spectrum therapeutics.

3.4 Conclusion

As demonstrated herein, the developed PepSAVI-MS pipeline is broad-spectrum, high-throughput, and has the potential to expedite the search for new bioactive peptides from plants. Our platform overcomes many traditional limitations of drug discovery efforts through achieving increased efficiency by directly targeting and characterizing only those species contributing to the bioactivity. This may eliminate many laborious rounds of bioassay-guided fractionation or the need for characterization of all molecular species present, regardless of demonstrated activity. Furthermore, the developed platform is highly versatile as it is adaptable to any natural product source of peptides and can test against diverse physiological targets, including bacteria, fungi, viruses, protozoans, and cancer cells for which there is a developed bioassay. With this platform there is little bias towards high abundance compounds as the trend of activity values is more important than the raw abundance values. While the PepSAVI-MS pipeline has been developed to target cationic AMPs, the platform is adaptable to target other types of compounds by changing the type of chromatography performed. The PepSAVI-MS pipeline opens the door

for investigating purpose-guided natural product extracts with a new lens, and has the potential to lead to the discovery of novel chemistries at the forefront of modern drug discovery.

3.5 Figures

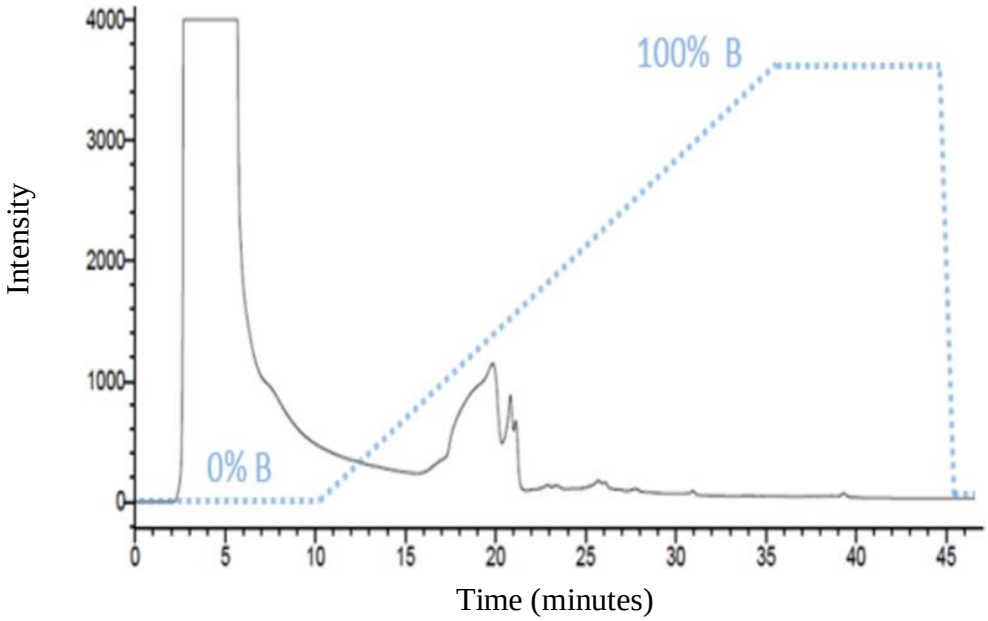


Figure 3.1 Strong cation exchange elution profile for *Viola odorata* (black) at 280 nm with the mobile phase gradient depicted with a blue dashed line.

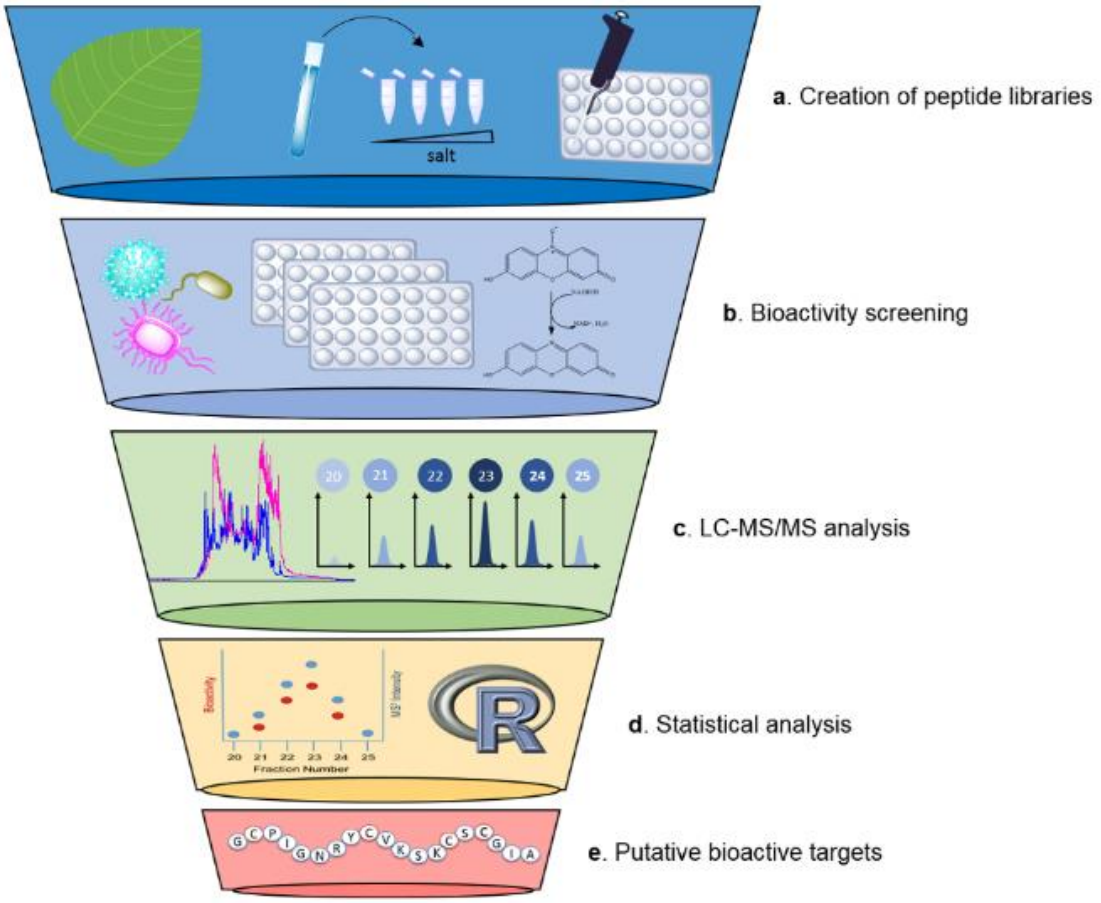


Figure 3.2 Overall workflow for the PepSAVI-MS screening platform including (a) creation of peptide libraries through extraction and fractionation of crude extracts, (b) whole-cell bioactivity screening of peptide libraries against pathogen targets of interest, (c) LC-MS/MS analysis of active peptide libraries and (d) statistical modeling of LC-MS/MS datasets vs. active bioactivity regions for determination of (e) putative bioactive peptide targets.

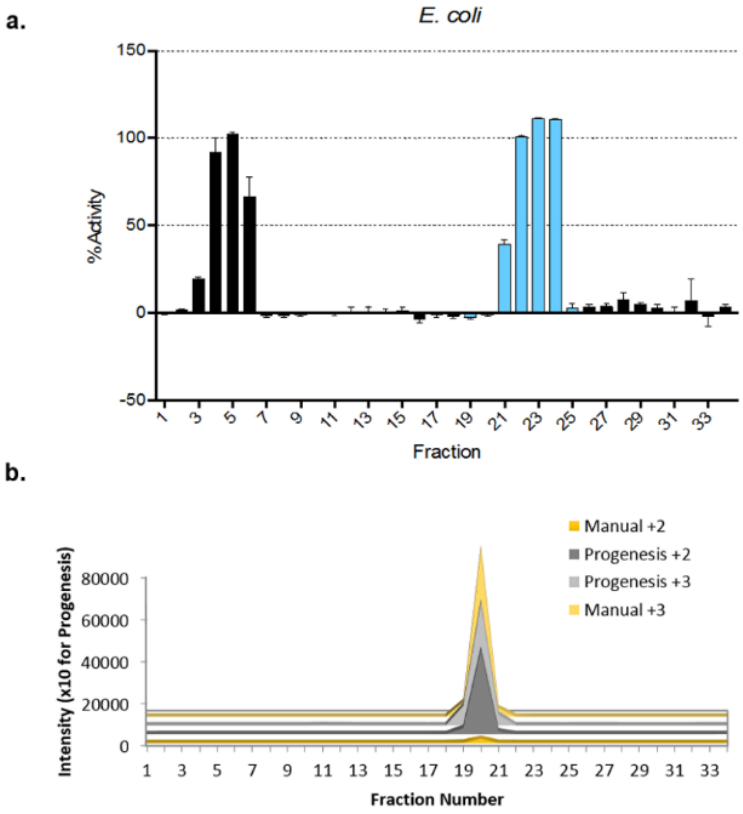


Figure 3.3 *V. odorata* fractions (a) bioactivity vs. *Escherichia coli* with the growth-inhibition defined bioactivity region in blue where % activity indicates the decrease in bacterial aerobic metabolism quantified by resazurin reduction. (b) Aligned cyO2 elution profile with comparison of manual (yellow) and automated (gray) extraction of the detected cyO2 charge states.

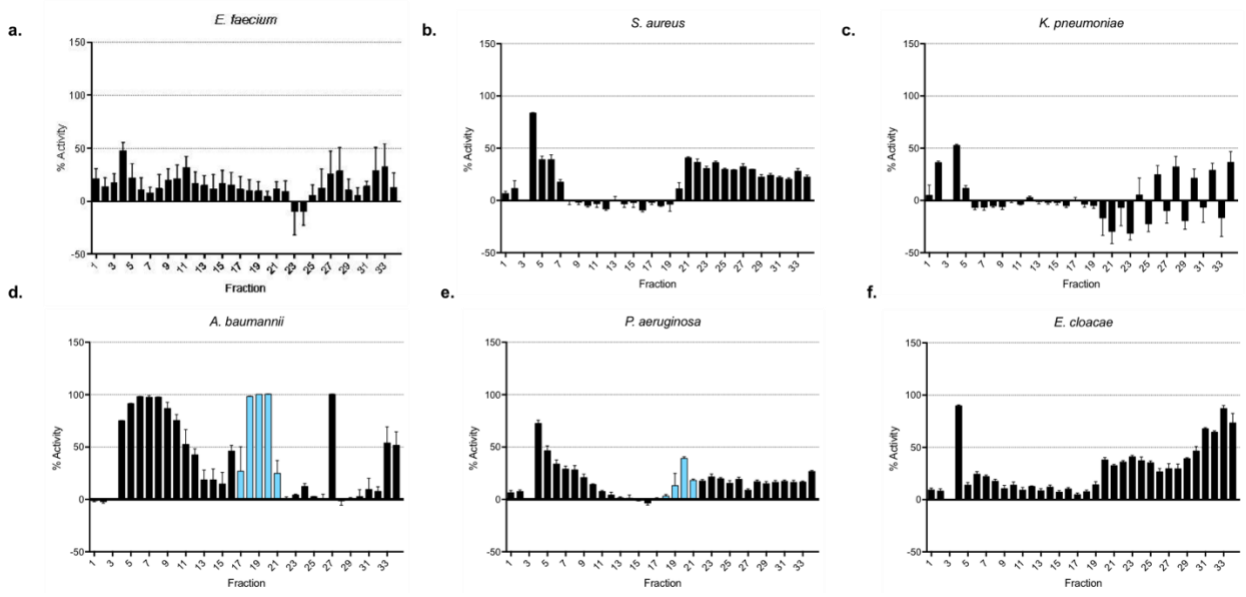


Figure 3.4 Bioactivity data representing the novel activity of *Viola odorata* fractions vs. ESKAPE pathogens: (a) *Enterococcus faecium* (b) *Staphylococcus aureus*, (c) *Klebsiella pneumoniae*, (d) *Acinetobacter baumannii*, (e) *Pseudomonas aeruginosa*, and (f) *Enterobacter cloacae*. The growth inhibition-defined bioactivity region of cyO2 is depicted using blue bars for the species deemed to demonstrate activity.

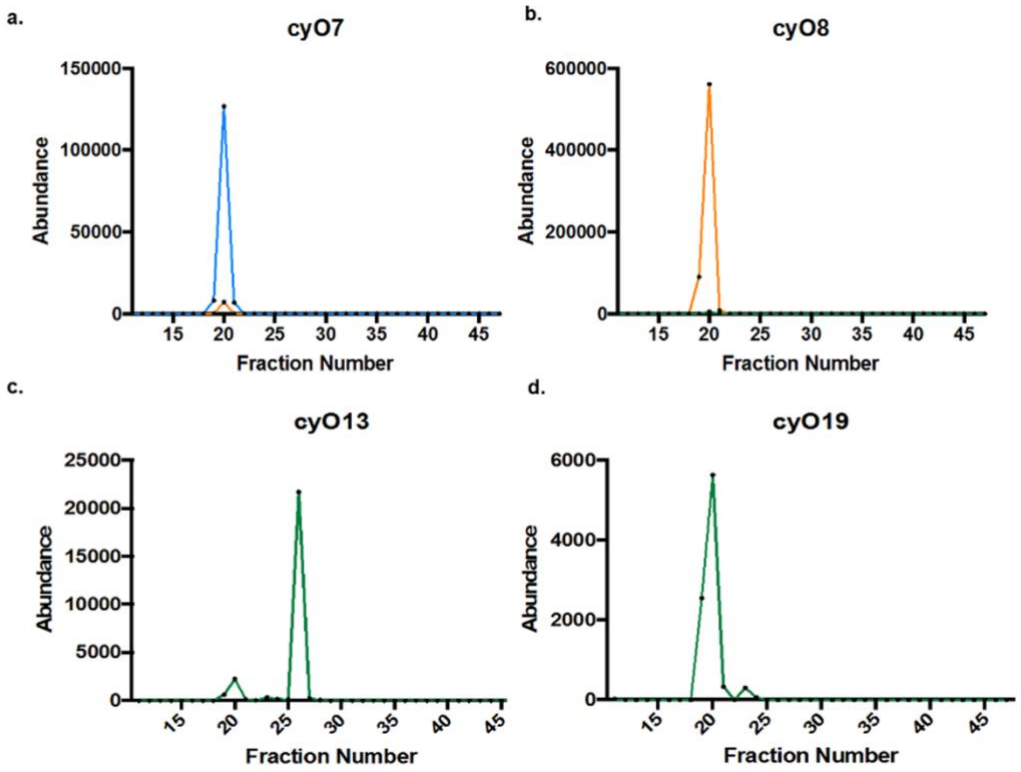


Figure 3.5 Strong cation exchange elution profiles of other detected cyclotides across *V. odorata* fractions 11 – 47, where the color of the line indicates the charge state(s) in which each cyclotide was detected: orange = +2, blue = +3, and green = +4.

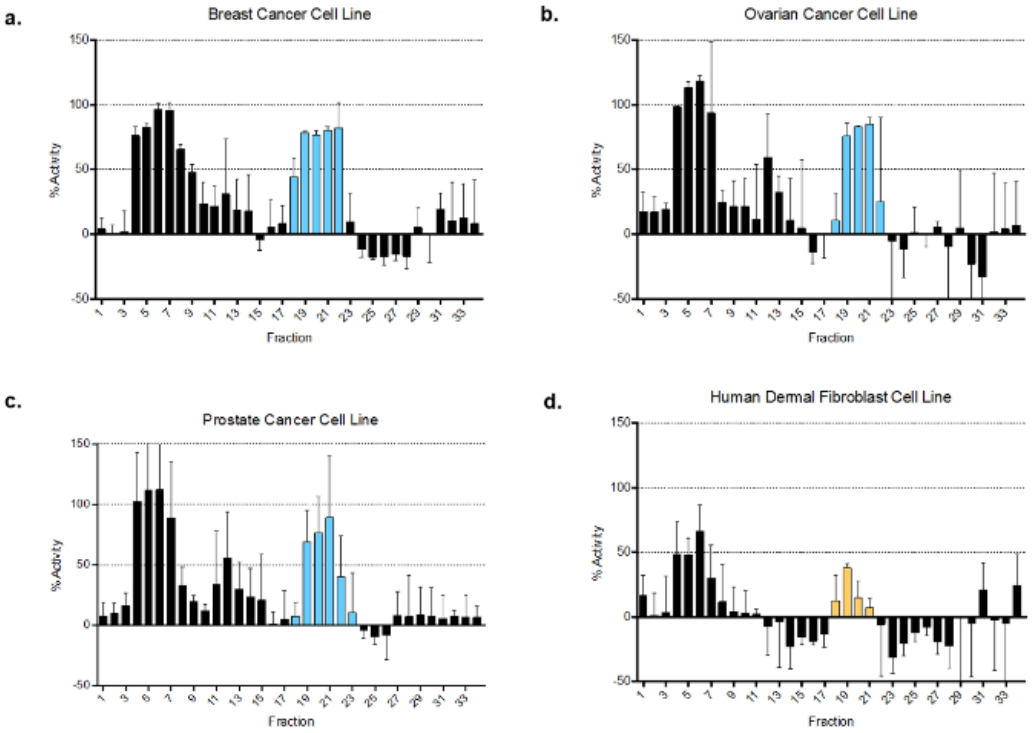


Figure 3.6 *V. odorata* peptide library bioactivity profile against (a) breast (MDA-MB-231), (b) ovarian (OVCAR-3), and (c) prostate (PC3) cancer cell lines and (d) a non-cancerous human dermal fibroblast cell line. The cytotoxicity-defined bioactivity region of cyO2 against the given cell line is defined with blue bars for cancer cell lines and yellow bars for the non-cancerous cell line.

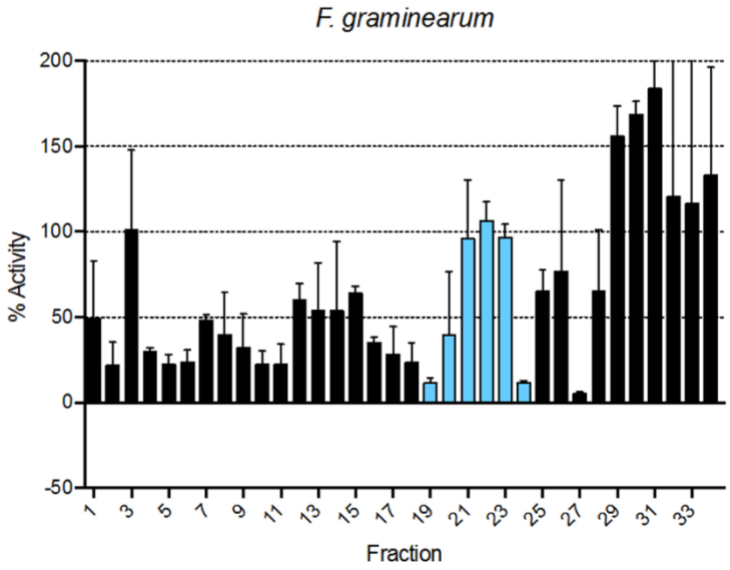


Figure 3.7 *V. odorata* peptide library bioactivity profile against the filamentous fungus *Fusarium graminearum* with the modeled bioactivity region depicted in blue.

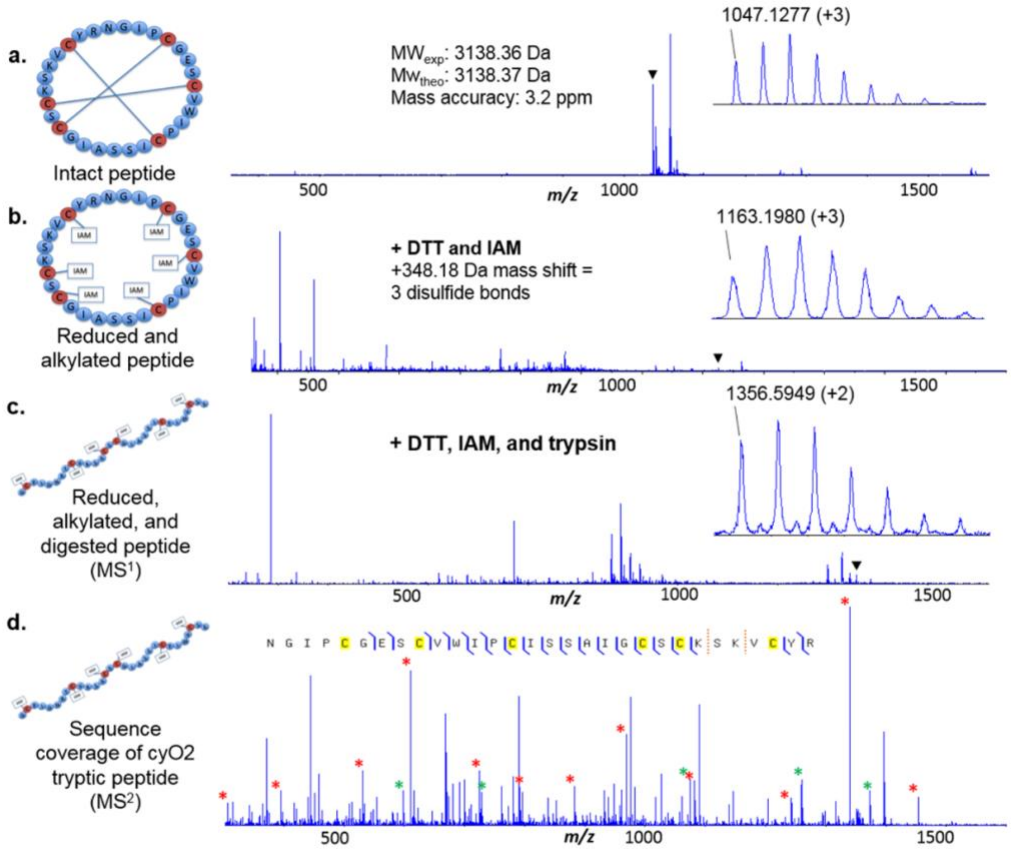


Figure 3.8 Mass spectral characterization of cyO2 including accurate mass spectrum of (a) intact peptide, (b) reduced and alkylated peptide, (c) tryptic peptide, and (d) MS/MS of tryptic peptide for sequence determination with inset graphical fragment map indicating b- (green) and y- (red) fragment ions (*) delineating primary sequence.

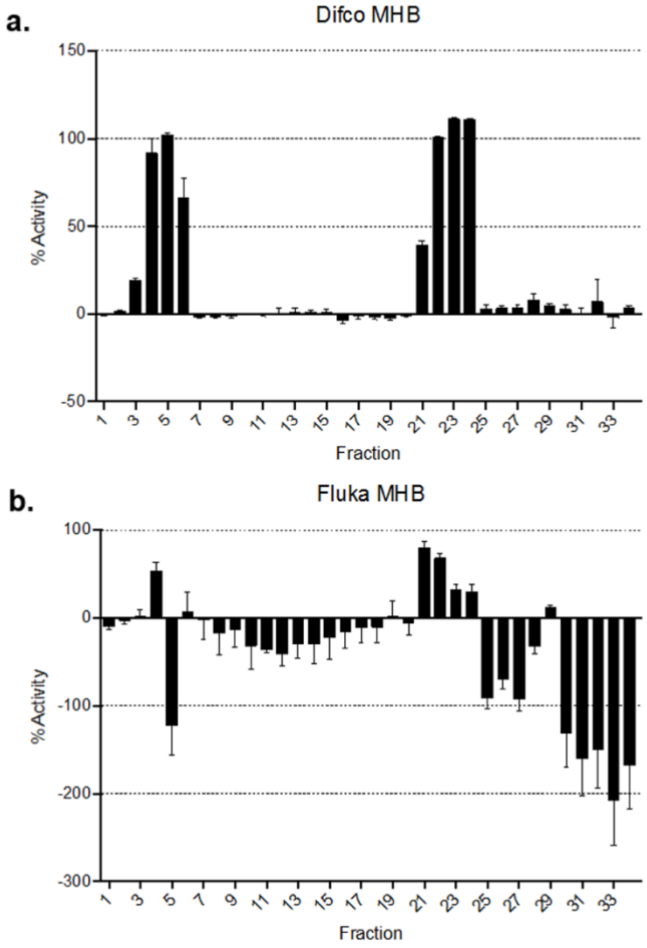


Figure 3.9 Difco MHB was used for all bacterial bioassays as shown with the *V. odorata* bioactivity against *E. coli* (a). When media from other vendors (e.g. Fluka) was used, bacterial growth was less robust, and some plant library fractions showed growth promotion indicating a nutritional deficiency in the MHB, supplied by those fractions, that interfered with accurate assessment of bioactivity (b). This effect was most pronounced with the Gram-positive organisms *E. faecalis* and *S. aureus*.

REFERENCES

1. Klevens, R. M.; Edwards, J. R.; Richards, C. L., Jr.; Horan, T. C.; Gaynes, R. P.; Pollock, D. A.; Cardo, D. M., Estimating health care-associated infections and deaths in U.S. hospitals, 2002. *Public Health Rep.* **2007,** *122* (2), 160-166.

2. Boucher, H. W.; Talbot, G. H.; Bradley, J. S.; Edwards, J. E.; Gilbert, D.; Rice, L. B.; Scheld, M.; Spellberg, B.; Bartlett, J., Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **2009,** *48* (1), 1-12.

3. Howard, C. R.; Fletcher, N. F., Emerging virus diseases: can we ever expect the unexpected? *Emerg Microbes Infect* **2012,** *1* (12), e46.

4. Guzman-Rodriguez, J. J.; Ochoa-Zarzosa, A.; Lopez-Gomez, R.; Lopez-Meza, J. E., Plant antimicrobial peptides as potential anticancer agents. *Biomed Res Int* **2015,** *2015*, 735087.

5. Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J., The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* **2015,** *14* (2), 111-129.

6. Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Goransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Muller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Sussmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **2013,** *30* (1), 108-160.

7. Essig, A.; Hofmann, D.; Munch, D.; Gayathri, S.; Kunzler, M.; Kallio, P. T.; Sahl, H. G.; Wider, G.; Schneider, T.; Aebi, M., Copsin, a novel peptide-based fungal antibiotic interfering with the peptidoglycan synthesis. *J. Biol. Chem.* **2014,** *289* (50), 34953-34964.

8. Schweizer, F., Cationic amphiphilic peptides with cancer-selective toxicity. *Eur. J. Pharmacol.* **2009,** *625* (1-3), 190-194.

9. Vlieghe, P.; Lisowski, V.; Martinez, J.; Khrestchatisky, M., Synthetic therapeutic peptides: science and market. *Drug Discov. Today* **2010,** *15* (1-2), 40-56.

10. Mandal, S. M.; Roy, A.; Ghosh, A. K.; Hazra, T. K.; Basak, A.; Franco, O. L., Challenges and future prospects of antibiotic therapy: from peptides to phages utilization. *Front. Pharmacol.* **2014,** *5*, (105), 1-12.

11. Walensky, L. D.; Bird, G. H., Hydrocarbon-Stapled Peptides: Principles, Practice, and Progress. *J. Med. Chem.* **2014,** *57* (15), 6275-6288.

12. Mohimani, H.; Kersten, R. D.; Liu, W. T.; Wang, M.; Purvine, S. O.; Wu, S.; Brewer, H. M.; Pasa-Tolic, L.; Bandeira, N.; Moore, B. S.; Pevzner, P. A.; Dorrestein, P. C., Automated genome mining of ribosomal peptide natural products. *ACS Chem. Biol.* **2014,** *9* (7), 1545-1551.

13. Henriques, S. T.; Huang, Y. H.; Castanho, M. A.; Bagatolli, L. A.; Sonza, S.; Tachedjian, G.; Daly, N. L.; Craik, D. J., Phosphatidylethanolamine binding is a conserved feature of cyclotide-membrane interactions. *J. Biol. Chem.* **2012,** *287* (40), 33629-33643.

14. Medema, M. H.; Paalvast, Y.; Nguyen, D. D.; Melnik, A.; Dorrestein, P. C.; Takano, E.; Breitling, R., Pep2Path: automated mass spectrometry-guided genome mining of peptidic natural products. *PLoS Comput. Biol.* **2014,** *10* (9), e1003822.

15. Skinnider, M. A.; Johnston, C. W.; Edgar, R. E.; Dejong, C. A.; Merwin, N. J.; Rees, P. N.; Magarvey, N. A., Genomic charting of ribosomally synthesized natural product chemical space facilitates targeted mining. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113* (42), E6343-E6351.

16. Kurita, K. L.; Glassey, E.; Linington, R. G., Integration of high-content screening and untargeted metabolomics for comprehensive functional annotation of natural product libraries. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** *112* (39), 11999-12004.

17. Fleeman, R.; LaVoi, T. M.; Santos, R. G.; Morales, A.; Nefzi, A.; Welmaker, G. S.; Medina-Franco, J. L.; Giulianotti, M. A.; Houghten, R. A.; Shaw, L. N., Combinatorial Libraries As a Tool for the Discovery of Novel, Broad-Spectrum Antibacterial Agents Targeting the ESKAPE Pathogens. *J. Med. Chem.* **2015,** *58* (8), 3340-3355.

18. Vizcaino, J. A.; Csordas, A.; del-Toro, N.; Dianes, J. A.; Griss, J.; Lavidas, I.; Mayer, G.; Perez-Riverol, Y.; Reisinger, F.; Ternent, T.; Xu, Q. W.; Wang, R.; Hermjakob, H., 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **2016,** *44* (D1), D447-456.

19. Herrmann, A.; Burman, R.; Mylne, J. S.; Karlsson, G.; Gullbo, J.; Craik, D. J.; Clark, R. J.; Göransson, U., The alpine violet, Viola biflora, is a rich source of cyclotides with potent cytotoxicity. *Phytochemistry* **2008,** *69* (4), 939-952.

20. Hammami, R.; Ben Hamida, J.; Vergoten, G.; Fliss, I., PhytAMP: a database dedicated to antimicrobial plant peptides. *Nucleic Acids Res.* **2009,** *37* (Database issue), D963-968.

21. Adochite, R. C.; Moshnikova, A.; Carlin, S. D.; Guerrieri, R. A.; Andreev, O. A.; Lewis, J. S.; Reshetnyak, Y. K., Targeting breast tumors with pH (low) insertion peptides. *Mol. Pharm.* **2014,** *11* (8), 2896-2905.

22. Garcia-Olmedo, F.; Molina, A.; Alamillo, J. M.; Rodriguez-Palenzuela, P., Plant defense peptides. *Biopolymers* **1998,** *47* (6), 479-491.

23. Epple, P.; Apel, K.; Bohlmann, H., An Arabidopsis thaliana thionin gene is inducible via a signal transduction pathway different from that for pathogenesis-related proteins. *Plant Physiol.* **1995,** *109* (3), 813-820.

24. Lee, S. C.; Hwang, B. K., CASAR82A, a pathogen-induced pepper SAR8.2, exhibits an antifungal activity and its overexpression enhances disease resistance and stress tolerance. *Plant Mol. Biol.* **2006,** *61* (1-2), 95-109.

25. van Loon, L. C.; Rep, M.; Pieterse, C. M., Significance of inducible defense-related proteins in infected plants. *Annu. Rev. Phytopathol.* **2006,** *44*, 135-162.

26. Sarker, S. D.; Nahar, L.; Kumarasamy, Y., Microtitre plate-based antibacterial assay incorporating resazurin as an indicator of cell growth, and its application in the in vitro antibacterial screening of phytochemicals. *Methods* **2007,** *42* (4), 321-324.

27. Zou, H.; Hastie, T., Regularization and variable selection via the elastic net. *J. Roy. Stat. Soc. Ser. B. (Stat. Method.)* **2005,** *67* (2), 301-320.

28. Zou, H.; Hastie, T. elasticnet: Elastic-Net for Sparse Estimation and Sparse PCA **2012**. http://CRAN.R-project.org/package=elasticnet.

29. Tibshirani, R., Regression Shrinkage and Selection via the Lasso. *J R Stat Soc Series B Stat Methodol.* **1996,** *58* (1), 267-288.

30. Efron, B.; Hastie, T.; Johnstone, I.; Tibshirani, R., Least angle regression. *Ann. Statist.* **2004,** *32* (2), 407-499.

31. Guerrero, A.; Lerno, L.; Barile, D.; Lebrilla, C. B., Top-down analysis of highly post-translationally modified peptides by Fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **2015,** *26* (3), 453-459.

32. Poth, A. G.; Colgrave, M. L.; Philip, R.; Kerenga, B.; Daly, N. L.; Anderson, M. A.; Craik, D. J., Discovery of cyclotides in the fabaceae plant family provides new insights into the cyclization, evolution, and distribution of circular proteins. *ACS Chem. Biol.* **2011,** *6* (4), 345-355.

33. Brownstein, N. C.; Guan, X.; Mao, Y.; Zhang, Q.; DiMaggio, P. A.; Xia, Q.; Zhang, L.; Marshall, A. G.; Young, N. L., Paired single residue-transposed Lys-N and Lys-C digestions for label-free identification of N-terminal and C-terminal MS/MS peptide product ions: ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry and tandem mass spectrometry for peptide de novo sequencing. *Rapid Commun. Mass Spectrom.* **2015,** *29* (7), 659-666.

34. Hayakawa, E.; Menschaert, G.; De Bock, P.-J.; Luyten, W.; Gevaert, K.; Baggerman, G.; Schoofs, L., Improving the Identification Rate of Endogenous Peptides Using Electron Transfer Dissociation and Collision-Induced Dissociation. *J. Proteome Res.* **2013,** *12* (12), 5410-5421.

35. Ni, W.; Lin, M.; Salinas, P.; Savickas, P.; Wu, S. L.; Karger, B. L., Complete mapping of a cystine knot and nested disulfides of recombinant human arylsulfatase A by multi-enzyme digestion and LC-MS analysis using CID and ETD. *J. Am. Soc. Mass. Spectrom.* **2013,** *24* (1), 125-133.

36. Ma, B.; Zhang, K.; Hendrie, C.; Liang, C.; Li, M.; Doherty-Kirby, A.; Lajoie, G., PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2003,** *17* (20), 2337-2342.

37. Spengler, B., De novo sequencing, peptide composition analysis, and composition-based sequencing: a new strategy employing accurate mass determination by fourier transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom* **2004,** *15* (5), 703-714.

38. Langsdorf, M.; Ghassempour, A.; Rompp, A.; Spengler, B., Characterization of a peptide family from the skin secretion of the Middle East tree frog Hyla savignyi by composition-based de novo sequencing. *Rapid Commun Mass Spectrom* **2010,** *24* (19), 2885-2899.

39. Ma, B.; Johnson, R., De novo sequencing and homology searching. *Molecular & cellular proteomics : MCP* **2012,** *11* (2), O111 014902.

40. Ireland, D. C.; Colgrave, M. L.; Craik, D. J., A novel suite of cyclotides from Viola odorata: sequence variation and the implications for structure, function and stability. *Biochem. J* **2006,** *400* (1), 1-12.

41. Pranting, M.; Loov, C.; Burman, R.; Goransson, U.; Andersson, D. I., The cyclotide cycloviolacin O2 from Viola odorata has potent bactericidal activity against Gram-negative bacteria. *J. Antimicrob. Chemother.* **2010,** *65* (9), 1964-1971.

42. Gerlach, S. L.; Rathinakumar, R.; Chakravarty, G.; Goransson, U.; Wimley, W. C.; Darwin, S. P.; Mondal, D., Anticancer and chemosensitizing abilities of cycloviolacin 02 from Viola odorata and psyle cyclotides from Psychotria leptothyrsa. *Biopolymers* **2010,** *94* (5), 617-625.

43. Lindholm, P.; Göransson, U.; Johansson, S.; Claeson, P.; Gullbo, J.; Larsson, R.; Bohlin, L.; Backlund, A., Cyclotides: a novel type of cytotoxic agents. *Mol. Cancer Ther.* **2002,** *1* (6), 365-369.

44. Göransson, U.; Sjogren, M.; Svangard, E.; Claeson, P.; Bohlin, L., Reversible antifouling effect of the cyclotide cycloviolacin O2 against barnacles. *J. Nat. Prod.* **2004,** *67* (8), 1287-1290.

45. Göransson, U.; Herrmann, A.; Burman, R.; Haugaard-Jönsson, L. M.; Rosengren, K. J., The Conserved Glu in the Cyclotide Cycloviolacin O2 Has a Key Structural Role. *ChemBioChem* **2009,** *10* (14), 2354-2360.

46. Lai, Y. C.; Yang, S. L.; Peng, H. L.; Chang, H. Y., Identification of genes present specifically in a virulent strain of Klebsiella pneumoniae. *Infect. Immun.* **2000,** *68* (12), 7149-7151.

47. Wu, K. M.; Li, L. H.; Yan, J. J.; Tsao, N.; Liao, T. L.; Tsai, H. C.; Fung, C. P.; Chen, H. J.; Liu, Y. M.; Wang, J. T.; Fang, C. T.; Chang, S. C.; Shu, H. Y.; Liu, T. T.; Chen, Y. T.; Shiau, Y. R.; Lauderdale, T. L.; Su, I. J.; Kirby, R.; Tsai, S. F., Genome sequencing and comparative analysis of Klebsiella pneumoniae NTUH-K2044, a strain causing liver abscess and meningitis. *J. Bacteriol.* **2009,** *191* (14), 4492-4501.

48. Park, I. Y.; Cho, J. H.; Kim, K. S.; Kim, Y. B.; Kim, M. S.; Kim, S. C., Helix stability confers salt resistance upon helical antimicrobial peptides. *J. Biol. Chem.* **2004,** *279* (14), 13896-13901.

49. Schwab, U.; Gilligan, P.; Jaynes, J.; Henke, D., In vitro activities of designed antimicrobial peptides against multidrug-resistant cystic fibrosis pathogens. *Antimicrob. Agents Chemother.* **1999,** *43* (6), 1435-1440.

CHAPTER 4: Fungal Secretome Analysis via PepSAVI-MS: Identification of the Bioactive Peptide KP4 from *Ustilago maydis*

*****Reprinted with permission, from Kirkpatrick, C.L., Parsley, N.C., Bartges, T.E., Cooke, M.E., Evans, W.S., Heil, L.R., Smith, T.J., Hicks, L.M. Fungal Secretome Analysis via PepSAVI-MS: Identification of the Bioactive Peptide KP4 from *Ustilago maydis*. JASMS **2018**.

4.1 Introduction

Fungi represent a diverse and versatile kingdom that has survived on earth for over 500 million years[1]. While ~100,000 fungal species have been described to date, it is estimated that 5 million fungal species could exist[2]. This vast and largely unexplored kingdom represents a rich source of untapped natural compounds that have been crucial for their survival. Used for defense, fungi produce a wide variety of molecules to leverage competitive advantage over invading microbial pathogens. Fungi-microbe interactions have led to the production of antimicrobial compounds that can be harnessed for human health, agricultural, and food safety applications[3-4]. Important among these are secreted small, cysteine rich antimicrobial peptides.

Recent discoveries increasingly reveal bioactive peptides from a variety of sources to have substantial structural diversity and broad applications[5-6]. The antimicrobial peptide, copsin, from the basidiomycete *Coprinopsis cinerea*, was first discovered in 2014 and has potent activity against gram-positive bacteria including both the human pathogen *Enterococcus faecium* and the foodborne pathogen *Listeria monocytogenes*[7]. Additionally, there has been a growing interest in the agricultural use of bioactive peptides for crop protection, including transgenic expression in plants and topical application as biopesticides. Spear-T, a bioinsecticide derived from spider venom and marketed by Vestaron, has recently been approved by the EPA for commercial use.

Importantly, unlike many currently used neonicotinoid-containing products, Spear-T has no adverse effects on bees or other beneficial insects, highlighting additional benefits of certain peptide-based agricultural products[8]. Fungi-microbe interactions have also inspired the use as these natural peptides in food safety applications, including the use of bacteriocins from lactic acid bacteria as starter cultures in food fermentation[9].

The promise of novel and effective mechanisms of action (MOAs) has revitalized peptide natural product discovery. In conjunction with methods aimed to expand knowledge of fungal genomes[10-11], methods for rapid AMP identification of these species are required. Current methods employed for AMP identification include bioassay-guided approaches which rely on multiple rounds of fractionation, require large amounts of material, and often bias toward highly abundant or highly active compounds. Alternative genome mining approaches leverage the advantage of economical deep sequencing technologies but require knowledge of antimicrobial gene clusters or amino acid sequences a priori and offer no direct measure of bioactivity. As such, efficient and versatile methods are needed to screen potential AMPs against common and emerging pathogens. We developed PepSAVI-MS (Statistically guided bioactive *pep*tides prioritized via mass spectrometry) to expedite discovery of bioactive peptides and validated this pipeline via identification and characterization of cycloviolacin O2 from the plant species *Viola odorata* (sweet violet)[12]. PepSAVI-MS utilizes selective extraction and fractionation of peptide source material, whole-cell bioactivity screening, and a statistics-guided mass spectrometry-based approach for targeted identification of only putatively bioactive compounds. To expand the search for potent and effective AMPs, we now extend this pipeline to fungal secretomes, a rich source of AMPs with potentially novel MOAs created and refined by extreme inter- and intra-species competition[3, 7, 13-14].

Herein, we demonstrate expansion of PepSAVI-MS to fungal-sourced AMPs using the killer toxin KP4 from *Ustilago maydis* P4 as proof-of-principle. KP4 is a highly positively charged 11.0 kDa peptide secreted by the corn smut fungus *U. maydis* when infected with the dsRNA P4 virus. The host and virus have co-evolved such that the host is not affected by the dsRNA virally-encoded toxin but when secreted by the host the toxin kills all other strains of *U. maydis*, thus giving the strain a selective advantage over other *U. maydis* strains[15-16]. Two adaptations were required to demonstrate applicability of PepSAVI-MS for fungal secretome analysis: 1) an approach to harvest peptides secreted into the growth media replaced the extraction procedure and 2) the bioassay format was adapted to screen against relevant fungal species. Successful application of PepSAVI-MS to microbial secretomes establishes the utility of this pipeline for novel fungal bioactive peptide discovery.

4.2 Materials and Methods

4.2.1 Fungal Strains and Growth Conditions

Ustilago maydis P4 and P6 strains were acquired from Robert Bozarth at Indiana State University and Jeremy Bruenn at Buffalo State, respectively. All *U. maydis* strains were grown in complete *U. maydis* media (UM media) consisting of 2.5% peptone (BD Difco), 1% dextrose (Sigma Aldrich), 0.15% ammonium nitrate (Sigma Aldrich), and 0.1% yeast extract (Sigma Aldrich). Four 5 mL starter cultures of *U. maydis* P4 were added to 2 L UM media and were grown to late-log phase at 25°C. Cells were removed by centrifugation (500 x g for 5 minutes) and the supernatant was collected.

4.2.2 Secretome Peptide Harvesting

The collected media was adjusted to a pH of 5.5 and stirred overnight with 140 mL of CM Sephadex C-25 resin (GE Healthcare) hydrated in 25 mM sodium acetate, pH 5.5. Slurry mixture

was gravity packed into a column and washed with two column volumes of 25 mM sodium acetate, pH 5.5, to remove unbound components. Peptides were eluted with 90 mL of 25 mM sodium acetate, pH 5.5, with 1 M NaCl, buffer exchanged into PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4), pH 7.3, using 3 kDa spin concentration filters (Millipore) and concentrated 10X.

4.2.3 Creation of Peptide Library via Crude SCX Fractionation

Concentrated peptide secretome sample (420 μL) was subjected to a 40-minute SCX method using a PolySulfoethyl A column (100 mm x 4.6 mm, 3 μm particles, PolyLC). A salt gradient was employed using a linear ramp of 5 mM ammonium formate, 20% acetonitrile, pH 2.7 to 500 mM ammonium formate, 20% acetonitrile, pH 3.0. Fractions were collected in one-minute increments and desalted with three washes of 1.3 mL deionized water using a vacuum concentrator.

4.2.4 Bioactivity Screening

The susceptible culture selected for *U. maydis* P4 killing assays was *U. maydis* P6 (UMP6) a related *Ustilago* strain infected with the P6 virus. While also producing a deadly toxin of its own, KP6, these cells are susceptible to KP4. UMP6 fungal challenge cultures were grown for 5 days in 5-mL aliquots of complete UM media at 25°C with 240 rpm shaking. Soft agar was prepared fresh on the fifth day by the addition of 1.5% bacto agar (BD) to complete UM media and cooled in a water bath to 48°C for 1 hour. UMP6 cultures were added to the cooled agar at the ratio of 5 mL fungus/250 mL agar. UMP6-infused agar was poured into 100 x 15 mL Petri dishes and once solidified, wells were carved into each plate using 10 μL pipette tips (4 wells/plate). UMP4 fractions (both crude and SCX) were added to each well (50 μL/fraction), in duplicate. After

compound addition, plates remained at room temperature until visible growth inhibition was present (~5 days). Radial zones of clearance were measured around the point of application.

4.2.5 LC-MS/MS Analysis of Peptide Library

The *U. maydis* P4 peptide library was analyzed via a nano-LC-ESI-MS/MS platform composed of a Waters nanoAcquity UPLC coupled to an AB Sciex TripleTOF 5600 QTOF mass spectrometer. Peptide fractions were diluted to ~0.2 μg/μL and acidified to 0.1 % formic acid. Five microliters of each sample were injected onto a trap column (NanoAcquity UPLC 2G-W/M Trap 5 μm Symmetry C18, 180 μm x 20 mm: Waters) before transfer to the analytical C18 column (10k PSI, 100 Å, 1.8 μm, 75 μm x 250 mm: Waters). Peptide separation was carried out at a flow rate of 0.3 μL/minute using a linear ramp of 5 – 50 % mobile phase B (mobile phase A, 0.1% formic acid; mobile phase B, 0.1% formic acid in acetonitrile) over 30 minutes. The MS was operated in positive ion, high sensitivity mode with the MS survey spectrum using a mass range of 350-1600 *m/z* in 250 ms and information dependent acquisition of MS/MS data, 87 ms per scan. For IDA MS/MS experiments, the first 20 features above 150 counts threshold and having a charge state of +2 to +5 were fragmented using rolling collision energy $\pm$5%. Each MS/MS experiment put the precursor *m/z* on an 8-second dynamic exclusion list. Auto calibration was performed every eight samples (8 h) to assure high mass accuracy in both MS and MS/MS acquisition. The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE [17] partner repository with the dataset identifier PXD006931. Deisotoped peak lists for each fraction were generated using Progenesis QI for Proteomics software (Nonlinear Dynamics, v.2.0). Automatic processing settings were used to align and peak pick ions across all runs. Identified features were quantified using AUC integration of survey scan data based on the summed intensity of each deisotoped feature. Data

was exported as “peptide ion data” with the default parameters from Progenesis at the “Identify Peptides” stage in the software.

4.2.6 Statistical Modeling

Areas of interest in the bioactivity profile must be selected for subsequent data reduction and modeling. The bioactivity region for *U. maydis* P4 vs. *U. maydis* P6 was defined based on the observed bioactivity profile as fractions 27 - 32. Using the PepSAVI-MS software package (https://cran.r-project.org/package=PepSAVIms)[12], background ions were eliminated through retention time (14 – 45 minutes), mass (2000 – 14000), and charge-state (2-10, inclusive) filters to reduce the data to potential compounds of interest. Retention time filters were selected to eliminate background ions, mass filters to select for the common mass range of bioactive peptides, and charge state filters to eliminate unwanted small molecules. Peak-picked data were binned and filtered using the previously established workflow-informed criteria[12]. Briefly, binning was performed with a 0.05 Da window of features with identical charge states and filtering required a maximum abundance inside the extended bioactivity area of interest (25 – 34) with <1% of that abundance outside of the chosen window. The extended bioactivity region was defined as the region of bioactivity $\pm$ 2 fractions on either side to account for the increased sensitivity of mass spectrometry (i.e. bioactive compounds may be present in those fractions at concentrations too low to detect bioactivity). All features required a minimum abundance of 1000. All *m/z* species meeting these filtering criteria were modeled using the elastic net estimator with a quadratic penalty parameter specification of 0.001 to determine each species’ contribution to the observed overall bioactivity profile. The resulting list contains candidate compounds ranked in order of when they entered the model, such that the highest ranked compounds have the greatest likelihood to be contributing to the bioactivity.

4.2.7 KP4 Purification

KP4 was purified as reported previously[18-20]. In brief, the toxin was isolated from the supernatant of the KP4 toxin expressing strains of *U. maydis* grown in complete *U. maydis* media for 7 to 10 days. Cells were removed by centrifugation at 10,000 x g for 30 minutes. The supernatant was stirred overnight with CM Sephadex C-25 resin (Amersham Biosciences) equilibrated with 25 mM sodium acetate, pH 5.5. The toxin was eluted with 1 M NaCl using a Pharmacia GradiFrac system. The eluent was concentrated using a Minitan II Ultrafiltration System (Millipore) with 1 kDa cutoff membranes and dialyzed against a 10 mM malonic acid, pH 6.0. KP4 was then purified using a high-resolution cation-exchange chromatography (Mono-S; Amersham Biosciences) matrix attached to a fast-performance liquid chromatography system in the same buffer and using sodium chloride for elution. The toxin was further purified with size exclusion chromatography using an Amersham Biosciences Superdex-75 gel filtration column. Toxin activity was tested throughout the purification using killing-zone activity assays and purity was assessed using Homogenous 20 SDS gels on an Amersham Biosciences Phastgel system. Only a single band representing KP4 was observable when silver staining was used to observe the protein bands.

4.2.8 KP4 Activity Validation

Purified KP4 was tested in the agar diffusion assay as described above. A dilution series of KP4 starting at 9 μM and decreasing two-fold to 0.018 μM was used to determine the minimum inhibitory concentration. Erythromycin was used as a positive control at 100 μg/mL. Radii of inhibition were measured 4 days after application to UMP6-infused agar. Killing assays at all KP4 and erythromycin concentrations were performed in duplicate.

4.3 Results and Discussion

4.3.1 Overview

PepSAVI-MS implements a multipronged approach for bioactive peptide discovery that utilizes selective isolation and fractionation of peptides from source material, bioactivity screening, and mass spectrometry-based peptidomics for the identification of putative bioactive peptide targets. As PepSAVI-MS was originally demonstrated for expressed plant peptides, two minor adaptations were required to demonstrate applicability for fungal secretome analysis: 1) an approach to harvest peptides secreted into the growth media replaced the extraction procedure and 2) the bioassay format was adapted to screen against relevant fungal species (Figure 4.1).

4.3.2 Creation of SCX Fractionated Libraries

Secreted peptides are collected from the cell-free supernatant using weak cation exchange resin added directly to the media. The peptide-resin slurry was gently washed to remove unretained compounds, and then high salt was used to elute peptides/proteins. Then, SCX fractionation was used to create a UMP4 peptide library for bioactivity screening.

4.3.3 Bioactivity Screening

PepSAVI-MS is amenable to any developed bioassay and can be modified to accommodate any target pathogen. High-throughput 96-well microtiter-based assays for bacterial species were presented in the original demonstration of PepSAVI-MS[12]. However, this bioassay format is not amenable to fungal species that often fail to grow to high density and form fungal mats that interfere with accurate bioactivity measurements. For fungal species, traditional diffusion assays offer a tried-and-true format to obtain bioactivity data[21]. As such, agar-based diffusion assays were used to examine the activity of the *U. maydis* P4 SCX fraction library against a susceptible *U. maydis* P6 culture, showing strong activity across a discrete and well-defined region (fractions

27-32) (Figure 4.2a). The crude nature of SCX allows for a single peptide to elute across multiple sequential fractions and thus span a wide bioactivity region. This characteristic is crucial for modeling of bioactivity and LC-MS data, as more points across the curve allow for increased confidence in matched profiles. The elution profile of a select subset of peptides across SCX fractions (obtained via LC-MS analysis) has been plotted in Figure 4.3. On average, any given peptide elutes across 2-3 sequential fractions in varying abundances, and peptides are eluting across all regions of the elution gradient. The combined elution and abundance profiles of each peptide create a unique fingerprint that allows for modeling with the bioactivity profile.

4.3.4 MS Profiling, Data Reduction, and Statistical Modeling

The *U. maydis* P4 SCX peptide library was subjected to LC-MS/MS analysis to obtain accurate mass and peptide abundance data for all constituents in the library. Peptide ion data reveals 17,473 features detected across the *U. maydis* P4 fraction set. Reduction of these features to those most-likely contributing to the bioactivity in each region narrowed the number of features to 19. This large reduction in features (99%) is unusual for what we typically see with other peptide libraries (~96% reduction in features[12] and unpublished data). However, examination of fractions containing KP4 showed minimal peptidyl complexity in these particular fractions. Because KP4 is highly positively charged (10 basic residues), it retained on the SCX column much longer than other peptides and was detected in late eluting fractions. Statistical modeling using elastic net penalized linear regression of this filtered data set will still be useful to reveal a candidate list of compounds likely responsible for the bioactivity. When applied to the *U. maydis* data set, the four detected KP4 charge states (Figure 4.2b-c) meeting the filtering criteria ranked in the top 19 contributors [+7 (1^{st} rank), +8 (5^{th}), +9 (6^{th}), +10 (4^{th})] to the bioactivity observed in fractions 27 – 32. Identification and ranking of multiple charge states for

a given compound imparts a built-in compound redundancy that can help prioritize ranked species for downstream characterization and activity validation. The rank order of each charge state was independent of respective abundance, indicating PepSAVI-MS does not bias for highly abundant species. While not done in this case, collapsing charge states via protein deconvolution prior to modeling would allow an opportunity for more unique compounds to rank highly.

4.3.5 Activity Validation with Purified KP4

To confirm the activity of KP4, purified peptide was used in the agar diffusion-based assay against UMP6. KP4 was added to each well using a dilution series from 9 - 0.018 μM for minimum inhibitory concentration determination. After 4 days of incubation with KP6 infused agar, zones of inhibition were measured for each KP4 concentration. As seen in Figure 4.4, KP4 displays potent, reproducible antifungal activity in a dose-dependent manner at concentrations as low as 0.07 μM. No visible growth inhibition was observed at the two lowest KP4 concentrations (35 and 18 nM). Interestingly, these results suggest that KP4 is more potent towards *U. maydis* P6 cells than towards *U. maydis* P2 cells (MIC: 0.33 μM[22]). KP4 is also reported to inhibit Arabidopsis root growth (ED_{50}: 8 μM[23]) and expressed calcium channels (IC_{50}:1.24 μM[16]) but to our knowledge, this is the lowest reported inhibitory activity of KP4.

Successful application of PepSAVI-MS for fungal species demonstrates rapid and accurate identification of bioactive components from complex natural product secretomes and allows for focused downstream characterization and validation experiments on only the most promising species. Expansion of PepSAVI-MS to this historically rich source of bioactive molecules greatly enhances its capabilities and potential for novel bioactive peptide discovery. Furthermore, future implementations of PepSAVI-MS in combination with workflows that yield accessibility of new scaffold products such as genome mining of intractable fungal species[10] and

strategies using advanced algorithms such as RiPPquest[24] or Pep2Path[25] leverage complementarity of each and continue to expand capabilities.

4.4 Conclusion

PepSAVI-MS is a versatile and easily adaptable pipeline for natural product bioactive peptide discovery. Minor modifications to peptide extraction/collection and bioactivity assay format facilitate exploration of fungal secretomes, a rich source of bioactive peptides. Successful application of PepSAVI-MS to microbial secretomes, as demonstrated with *U. maydis*, opens the door to investigating fungal species with a new lens and may have translational applications ranging from human health, food safety, and agriculture.

4.5 Figures

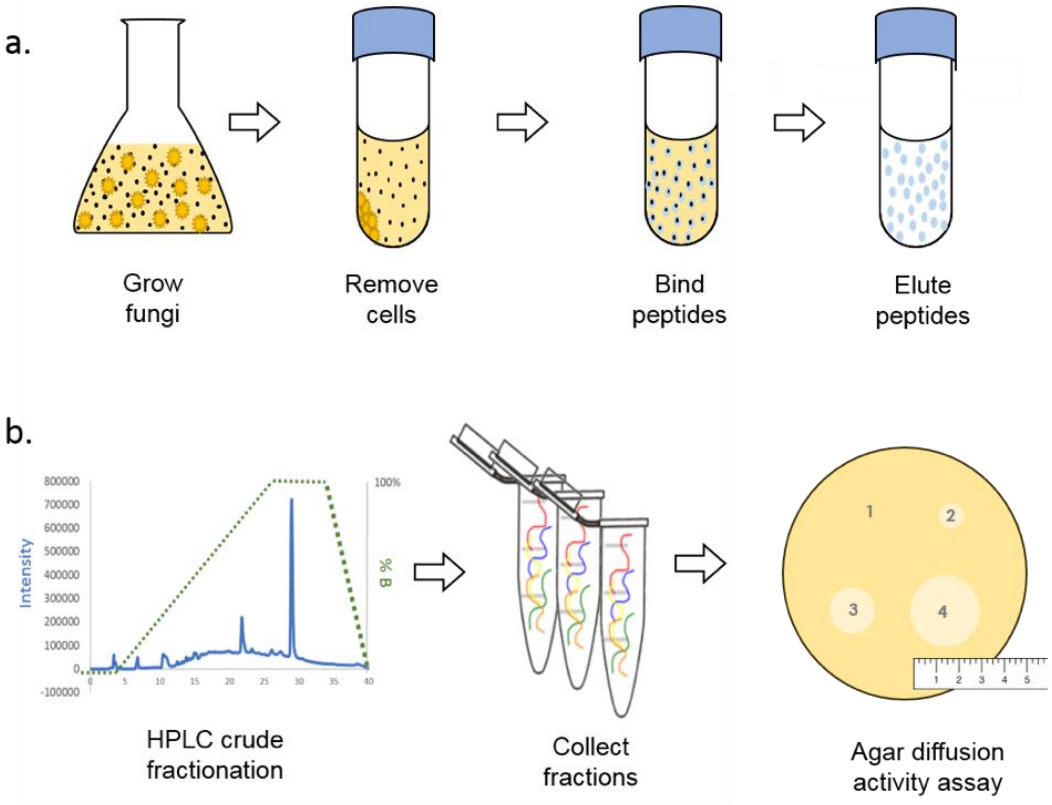


Figure 4.1 Modifications to the originally proposed PepSAVI-MS pipeline for (a) collection of secreted peptides and (b) creation of fraction libraries and addition of each library to the adapted bioassay screen against relevant fungal species. Harvesting of secreted peptides includes large-culture microbial growth, secretome isolation via centrifugation, addition of weak cation exchange peptide binding resin to cell-free supernatent, and elution of peptides off of the resin. Eluted peptides are then fractionated into libraries that are subject to bioactivity screening via agar diffusion activity assays.

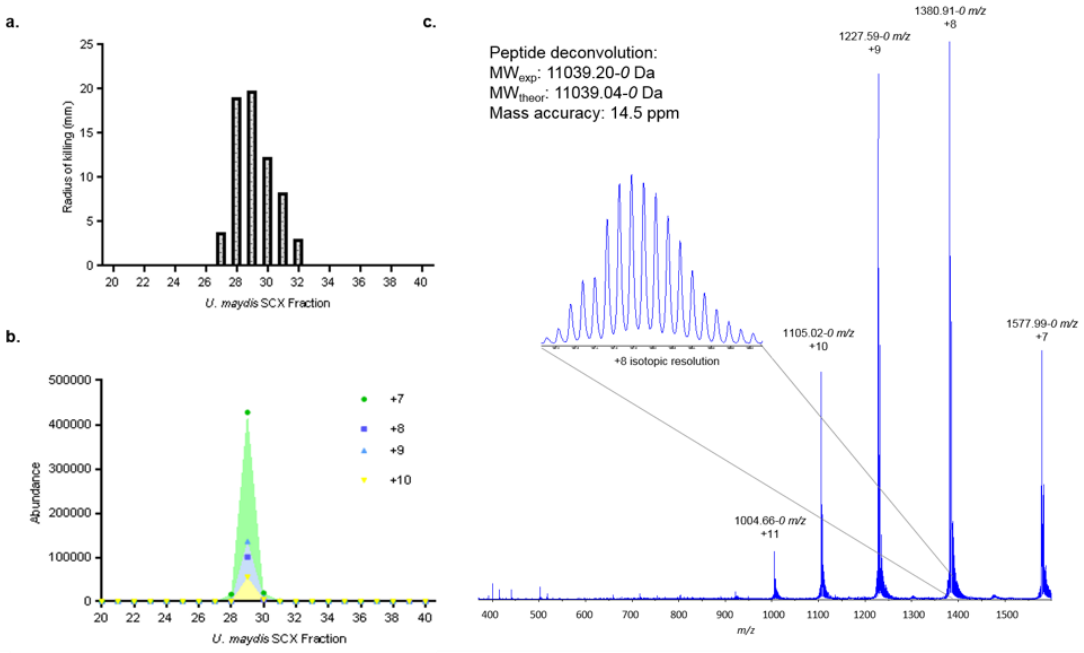


Figure 4.2 *U. maydis* bioactivity and KP4 LC-MS analysis. (a) *U. maydis* fraction library bioactivity screening against UMP6 with (b) aligned elution profile for KP4 as determined via manual charge state extraction. (c) Mass spectrum of intact KP4.

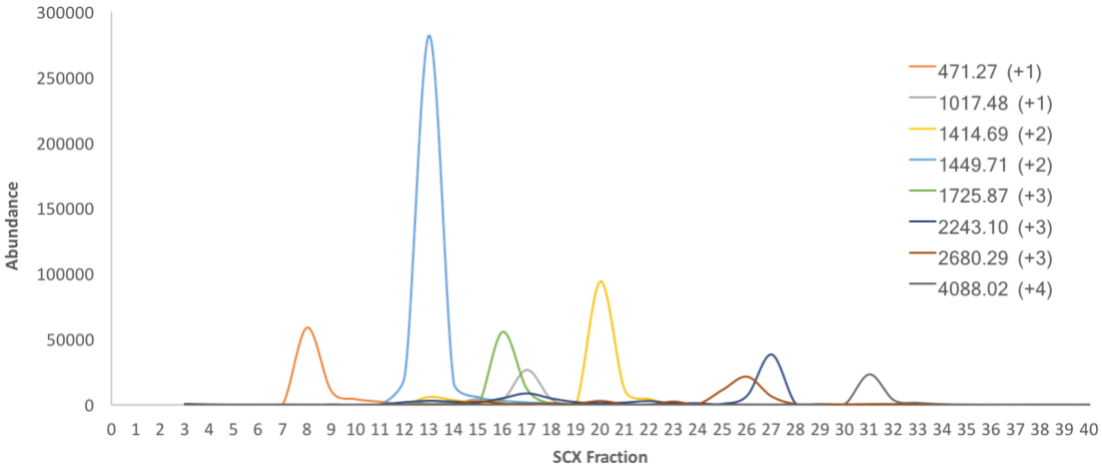


Figure 4.3 Elution profiles for a select subset of peptides (denoted by their mass and detected charge state) generated from exported peptide ion data. Distributions are plotted for a range of peptides across SCX fractions with relatively similar abundances for observation on the same scale; however, lower and higher abundance species display similar elution profiles. Representative peptides elute across the entire SCX gradient with an average peak width of 2-3 fractions.

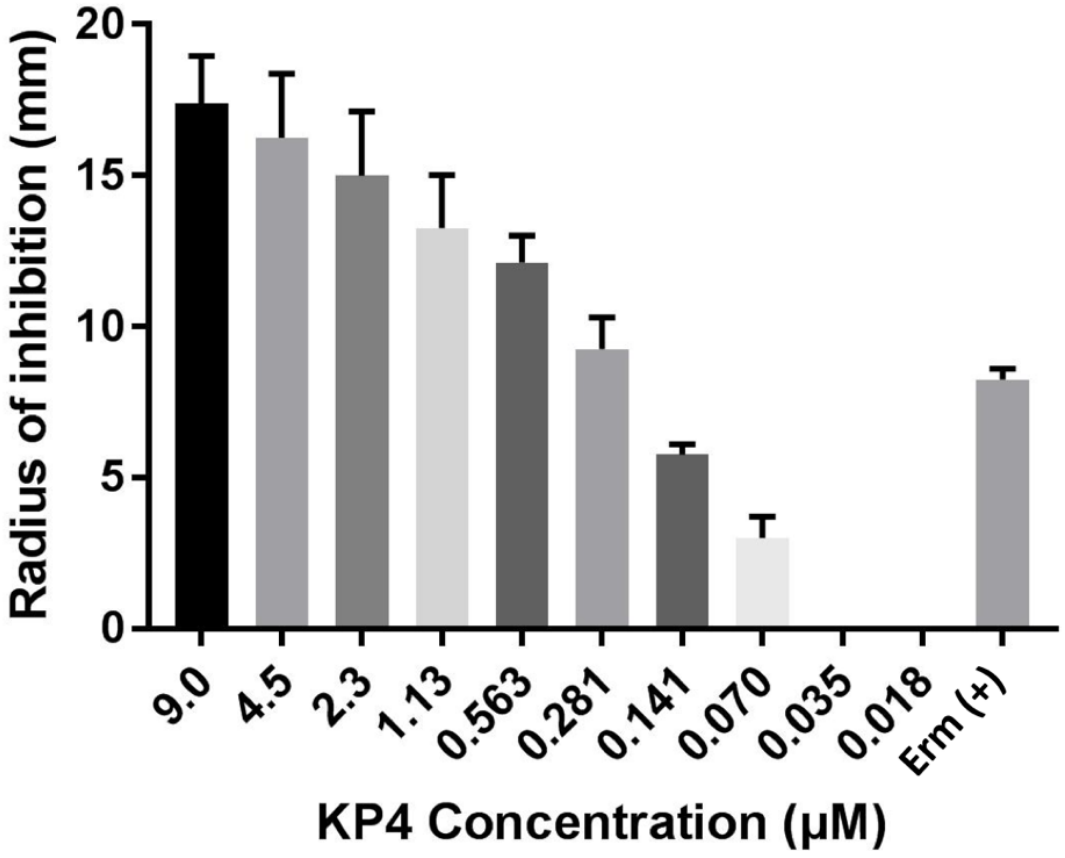


Figure 4.4 Radii of inhibition observed through agar diffusion assays testing purified KP4 against the challenge culture U. maydis P6. KP4 was added to each well in varying concentrations using a dilution series of KP4 starting at 9 μM and decreasing two-fold to 0.018 μM. Erythromycin (Erm (+)) was used as a positive control at 100 μg/mL. Zones of inhibition were measured 4 days after application to UMP6-infused agar. Killing assays at all KP4 and Erm concentrations were performed in duplicate.

REFERENCES

1. Brundrett, M. C., Coevolution of roots and mycorrhizas of land plants. *New Phytol.* **2002,** *154* (2), 275-304.

2. Blackwell, M., The fungi: 1, 2, 3 ... 5.1 million species? *Am. J. Bot.* **2011,** *98* (3), 426-438.

3. Scherlach, K.; Graupner, K.; Hertweck, C., Molecular bacteria-fungi interactions: effects on environment, food, and medicine. *Annu. Rev. Microbiol.* **2013,** *67*, 375-397.

4. Dang, T.; Süssmuth, R. D., Bioactive Peptide Natural Products as Lead Structures for Medicinal Use. *Acc. Chem. Res.* **2017,** *50* (7), 1566-1576.

5. Mandal, S. M.; Roy, A.; Ghosh, A. K.; Hazra, T. K.; Basak, A.; Franco, O. L., Challenges and future prospects of antibiotic therapy: from peptides to phages utilization. *Front. Pharmacol.* **2014,** *5* (105), 1-12.

6. Walensky, L. D.; Bird, G. H., Hydrocarbon-stapled peptides: principles, practice, and progress. *J. Med. Chem.* **2014,** *57* (15), 6275-6288.

7. Essig, A.; Hofmann, D.; Munch, D.; Gayathri, S.; Kunzler, M.; Kallio, P. T.; Sahl, H. G.; Wider, G.; Schneider, T.; Aebi, M., Copsin, a novel peptide-based fungal antibiotic interfering with the peptidoglycan synthesis. *J. Biol. Chem.* **2014,** *289* (50), 34953-34964.

8. Bomgardner, M. M., Spider venom: An insecticide whose time has come? *Chem. Eng. News* **2017,** *95* (11), 30-31.

9. Golneshin, A.; Adetutu, E.; Ball, A. S.; May, B. K.; Van, T. T.; Smith, A. T., Complete Genome Sequence of Lactobacillus plantarum Strain B21, a Bacteriocin-Producing Strain Isolated from Vietnamese Fermented Sausage Nem Chua. *Genome Announc* **2015,** *3* (2), e00055-00015.

10. Clevenger, K. D.; Bok, J. W.; Ye, R.; Miley, G. P.; Verdan, M. H.; Velk, T.; Chen, C.; Yang, K.; Robey, M. T.; Gao, P.; Lamprecht, M.; Thomas, P. M.; Islam, M. N.; Palmer, J. M.; Wu, C. C.; Keller, N. P.; Kelleher, N. L., A scalable platform to identify fungal secondary metabolites and their gene clusters. *Nat. Chem. Biol.* **2017,** *13* (8), 895-01.

11. Grigoriev, I. V.; Cullen, D.; Goodwin, S. B.; Hibbett, D.; Jeffries, T. W.; Kubicek, C. P.; Kuske, C.; Magnuson, J. K.; Martin, F.; Spatafora, J. W.; Tsang, A.; Baker, S. E., Fueling the future with fungal genomics. *Mycology* **2011,** *2* (3), 192-209.

12. Kirkpatrick, C. L.; Broberg, C. A.; McCool, E. N.; Lee, W. J.; Chao, A.; McConnell, E. W.; Pritchard, D. A.; Hebert, M.; Fleeman, R.; Adams, J.; Jamil, A.; Madera, L.; Stromstedt, A. A.; Goransson, U.; Liu, Y.; Hoskin, D. W.; Shaw, L. N.; Hicks, L. M., The "PepSAVI-MS" Pipeline for Natural Product Bioactive Peptide Discovery. *Anal. Chem.* **2017,** *89* (2), 1194-1201.

13. Amaral, A. C.; Silva, O. N.; Mundim, N. C. C. R.; de Carvalho, M. J. A.; Migliolo, L.; Leite, J. R. S. A.; Prates, M. V.; Bocca, A. L.; Franco, O. L.; Felipe, M. S. S., Predicting antimicrobial peptides from eukaryotic genomes: In silico strategies to develop antibiotics. *Peptides* **2012,** *37* (2), 301-308.

14. Mygind, P. H.; Fischer, R. L.; Schnorr, K. M.; Hansen, M. T.; Sonksen, C. P.; Ludvigsen, S.; Raventos, D.; Buskov, S.; Christensen, B.; De Maria, L.; Taboureau, O.; Yaver, D.; Elvig-Jorgensen, S. G.; Sorensen, M. V.; Christensen, B. E.; Kjaerulff, S.; Frimodt-Moller, N.; Lehrer, R. I.; Zasloff, M.; Kristensen, H. H., Plectasin is a peptide antibiotic with therapeutic potential from a saprophytic fungus. *Nature* **2005,** *437* (7061), 975-980.

15. Allen, A.; Islamovic, E.; Kaur, J.; Gold, S.; Shah, D.; Smith, T. J., Transgenic maize plants expressing the Totivirus antifungal protein, KP4, are highly resistant to corn smut. *Plant Biotechnol. J.* **2011,** *9* (8), 857-864.

16. Gage, M. J.; Rane, S. G.; Hockerman, G. H.; Smith, T. J., The virally encoded fungal toxin KP4 specifically blocks L-type voltage-gated calcium channels. *Mol. Pharmacol.* **2002,** *61* (4), 936-944.

17. Vizcaino, J. A.; Csordas, A.; del-Toro, N.; Dianes, J. A.; Griss, J.; Lavidas, I.; Mayer, G.; Perez-Riverol, Y.; Reisinger, F.; Ternent, T.; Xu, Q. W.; Wang, R.; Hermjakob, H., 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **2016,** *44* (D1), D447-456.

18. Gu, F.; Khimani, A.; Rane, S. G.; Flurkey, W. H.; Bozarth, R. F.; Smith, T. J., Structure and function of a virally encoded fungal toxin from Ustilago maydis: a fungal and mammalian Ca2+ channel inhibitor. *Structure* **1995,** *3* (8), 805-814.

19. Gu, F.; Khimani, A.; Rane, S.; Flurkey, W. H.; Bozarth, R. F.; Smith, T. J., The Structure of a Virally Encoded Fungal Toxin from Ustilago Maydis that Inhibits Fungal and Mammalian Calcium Channels. In *Protein Toxin Structure*, Springer Berlin Heidelberg: Berlin, Heidelberg, **1996**, 291-303.

20. Gu, F.; Sullivan, T. S.; Che, Z.; Ganesa, C.; Flurkey, W. H.; Bozarth, R. F.; Smith, T. J., The characterization and crystallization of a virally encoded Ustilago maydis KP4 toxin. *J. Mol. Biol.* **1994,** *243* (4), 792-795.

21. Tao, J.; Ginsberg, I.; Banerjee, N.; Held, W.; Koltin, Y.; Bruenn, J. A., Ustilago maydis KP6 killer toxin: structure, expression in Saccharomyces cerevisiae, and relationship to other cellular toxins. *Mol. Cell. Biol.* **1990,** *10* (4), 1373-1381.

22. Gage, M. J.; Bruenn, J.; Fischer, M.; Sanders, D.; Smith, T. J., KP4 fungal toxin inhibits growth in Ustilago maydis by blocking calcium uptake. *Mol. Microbiol.* **2001,** *41* (4), 775-785.

23. Allen, A.; Snyder, A. K.; Preuss, M.; Nielsen, E. E.; Shah, D. M.; Smith, T. J., Plant defensins and virally encoded fungal toxin KP4 inhibit plant root growth. *Planta* **2008,** *227* (2), 331-339.

24. Mohimani, H.; Kersten, R. D.; Liu, W. T.; Wang, M.; Purvine, S. O.; Wu, S.; Brewer, H. M.; Pasa-Tolic, L.; Bandeira, N.; Moore, B. S.; Pevzner, P. A.; Dorrestein, P. C., Automated genome mining of ribosomal peptide natural products. *ACS Chem. Biol.* **2014,** *9* (7), 1545-1551.

25. Medema, M. H.; Paalvast, Y.; Nguyen, D. D.; Melnik, A.; Dorrestein, P. C.; Takano, E.; Breitling, R., Pep2Path: automated mass spectrometry-guided genome mining of peptidic natural products. *PLoS Comput. Biol.* **2014,** *10* (9), e1003822.

CHAPTER 5: Exploring Bioactive Peptides from Bacterial Secretomes Using PepSAVI-MS: Identification and Characterization of Bac-21 from *Enterococcus faecalis* pPD1

*****Manuscript submitted for publication. Authors: Kirkpatrick, C.L., Parsley, N.C., Bartges, T.E., Wing, C.E., Kommineni S., Kristich C.J., Salzman, N.H., Patrie, S.M., and Hicks, L.M.

5.1 Introduction

Antibiotic-resistant infections are widespread, constantly evolving, and one of the biggest threats to global health (World Health Organization, 2017). While multidrug resistance (MDR) is occurring across an alarming number of bacterial species, among the most threatening MDR gram-positive organisms are enterococci. Though common commensal bacteria, enterococci are leading causes of hospital-acquired infections and are becoming highly antibiotic resistant[1-2]. Many species display intrinsic resistance to beta-lactam antibiotics, and further acquired resistance to tetracyclines, rifampin, and vancomycin, among others, severely limits therapeutic treatment options[2-3]. While the *Enterococcus* genus includes over 17 species, only a few have been shown to cause infections in humans. *Enterococcus faecalis* and *Enterococcus faecium* are the two most prevalent infectious species in humans and account for >90% of human isolates[4]. It is difficult, if not impossible in some cases, to treat these emerging and highly resistant enterococcal strains. It is crucial to develop analytical tools and novel therapeutic strategies to both characterize and treat these deadly strains.

Many lactic acid bacteria, including enterococci, produce bacteriocins - small cyclic peptides capable of killing or inhibiting growth of related or similar species[5-8]. Bacteriocin production is one tool bacteria can use to gain competitive advantage in an environment and establish a stable

niche for the producing strain[5, 9-10]. The enterococcal bacteriocin Bac-21 is encoded for on the sex pheromone-responsive conjugative plasmid pPD1[11]. This class of plasmids constitutes a complex and highly efficient system used to both transfer genetic information and function as a regulator of bacterial virulence. To date, over 20 of these plasmids have been identified but only among enterococcal species[12]. The pPD1 plasmid is predicted to produce a mature, 70 amino acid bacteriocin, Bac-21, identical in nucleotide sequence to another well-characterized bacteriocin, AS-48[11, 13]. However, the active Bac-21 functional peptide has yet to be characterized. As demonstrated recently, the Bac-21 producing enterococcal strain is able to specifically decolonize antibiotic-resistant enterococci from the gastrointestinal tract[14]. This has major implications as a novel therapeutic strategy to overcome difficult-to-treat enterococcal infections, and by extension a similar approach could be used to combat other problematic gut organisms. While this highlights potential therapeutic implementation of bacteriocins, we posit that these peptide toxins represent a large and relatively untapped source of bioactive peptides that can be harnessed for drug discovery applications.

PepSAVI-MS (Statistically guided bioactive *pep*tides prioritized via mass spectrometry) was developed to expedite the search for botanical bioactive peptides while addressing limitations of bioassay-guided fractionation and genome mining approaches[15]. To enhance the search for potent and effective antimicrobial peptides (AMPs), we now extend this pipeline to bacterial secretomes, a rich source of AMPs with potentially novel mechanisms of action (MOAs) created and refined by extreme inter- and intra-species competition[14, 16-17]. Herein, we demonstrate expansion of PepSAVI-MS to bacterial-sourced AMPs using the bacteriocin Bac-21 from *Enterococcus faecalis* harboring the pPD1 plasmid. A robust agar diffusion assay approach was implemented to screen bioactive peptides against infectious enterococcal strains. Additionally,

we experimentally validate the Bac-21 peptide sequence and post-translational processing for the first time. Successful application of PepSAVI-MS to bacterial secretomes is demonstrated with this representative species and establishes proof-of-principle for use in novel microbial bioactive peptide discovery.

5.2 Materials and Methods

5.2.1 Microbial Strains and Growth Conditions

E. faecalis (CK135) harboring pPD1[14, 18-19] and *E. faecium* (JL282)[18] strains were described previously. All *Enterococcus* strains were grown in Mueller Hinton Broth (MHB, BD Difco) and four 5 mL starter cultures of *E. faecalis* were added to 2 L of MHB and were grown to late-log phase at 37°C. Cells were removed by centrifugation (500 x g for 5 minutes) and the supernatant was collected.

5.2.2 Creation of Peptide Libraries

Secretome capture. The cell-free supernatant was adjusted to a pH of 5.5 and stirred overnight with 140 mL of CM Sephadex C-25 resin hydrated in 25 mM sodium acetate, pH 5.5. Slurry mixture was gravity packed into a column and washed with two column volumes of 25 mM sodium acetate, pH 5.5, to remove unbound components. Peptides were eluted with 25 mM sodium acetate, pH 5.5, with 1 M NaCl and five 50 mL fractions were collected. Fractions were buffer exchanged into 5 mM ammonium formate, pH 2.7, using 3 kDa spin concentration filters (Millipore) and concentrated 10X.

Strong cation exchange fractionation. Enterococcus crude fractions 2 and 3 (250 μL/fraction) were combined for strong cation exchange (SCX) fractionation via HPLC. The sample was subject to a 55-minute SCX method using a PolySulfoethyl A column (100 mm x 4.6 mm, 3 μm

particles, PolyLC). A salt gradient was employed using a linear ramp of 5 mM ammonium formate, 20% acetonitrile, pH 2.7 to 500 mM ammonium formate, 20% acetonitrile, pH 3.0. Fractions were collected in one-minute increments and desalted with three washes of 1.3 mL deionized water using a vacuum concentrator.

5.2.3 Bioactivity Screening

The susceptible culture selected for *E. faecalis* pPD1 killing assays, *E. faecium* JL282, was grown overnight in MHB. Soft agar was prepared for the agar diffusion assay by the addition of 1.5% bacto agar (BD) to broth solutions. Microbes were added to warm (48°C), soft agar and poured into 100 x 15 mm culture dishes. Once solidified, wells were carved in the agar and 50 μL of each test fraction was added to a different well. Plates were incubated at 37°C until visible growth inhibition was present (~2 days). Radial zones of clearance were measured around the point of application. Bacterial assays were performed in duplicate.

5.2.4 LC-MS/MS Analysis of Digested Peptide Library

The *E. faecalis* peptide library was subject to reduction, alkylation, and Glu-C digestion (*Staphylococcus aureus* Protease V8) prior to LC-MS/MS analysis. Fractions were reduced using 10 mM dithiothreitol (30 minutes, 45°C, 850 rpm) and subsequently alkylated with 100 mM iodoacetamide (15 minutes, 25°C, 850 rpm) prior to overnight digestion with Glu-C (37°C, 850 rpm). *E. faecalis* digested fractions were cleaned up with Pierce C18 zip tips (Thermo Fisher Scientific) before subsequent LC-MS analysis. Peptide libraries were analyzed via a nano-LC-ESI-MS/MS platform: Waters nanoAcquity UPLC coupled to an AB Sciex TripleTOF 5600. Peptide fractions were diluted to ~0.2 μg/μL and acidified to 0.1 % formic acid. Five microliters of each sample were injected onto a trap column (NanoAcquity UPLC 2G-W/M Trap 5 μm Symmetry C18, 180 μm x 20 mm: Waters) before transfer to the analytical C18 column (10k

PSI, 100 Å, 1.8 µm, 75 µm x 250 mm: Waters). Peptide separation was carried out at a flow rate of 0.3 µL/minute using a linear ramp of 5 – 50 % B (mobile phase A, 0.1% formic acid; mobile phase B, 0.1% formic acid in acetonitrile) over 30 minutes. The MS was operated in positive ion, high sensitivity mode with the MS survey spectrum using a mass range of 350-1600 *m/z* in 250 ms and information dependent acquisition (IDA) of MS/MS data, 87 ms per scan. For IDA MS/MS experiments, the first 20 features above 150 counts threshold and having a charge state of +2 to +5 were fragmented using rolling collision energy ±5%. Each MS/MS experiment put the precursor *m/z* on an 8-second dynamic exclusion list. Auto calibration was performed every eight samples (8 h) to assure high mass accuracy in both MS and MS/MS acquisition. The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE[20] partner repository with the dataset identifier PXD009003 and 10.6019/PDX009003. Deisotoped peak lists for each fraction were generated using Progenesis QI for Proteomics software (Nonlinear Dynamics, v.2.0). Automatic processing settings were used to align all runs using fraction 37 as the reference and peak picking ions across the entire digested library. Identified features were quantified using AUC integration of survey scan data based on the summed intensity of each deisotoped feature. Data was exported as "peptide ion data" with the default parameters from Progenesis at the "Identify Peptides" stage in the software.

5.2.5 Database Searching of Digested Peptide Library

Identification and location of Bac-21 Glu-C digested peptides was determined using Mascot (v.2.5.0; Matrix Science, http://www.matrixscience.com/). While the mature Bac-21 peptide has until now not been physically detected or molecularly characterized, its identity in nucleotide sequence to another well-characterized bacteriocin, AS-48, was used to predict its protein sequence. Database searching was performed using the *Firmicutes* taxonomy of the SwissProt

database (68,530 entries; accessed February, 2017) appended with the predicted peptide sequence of Bac-21. Searches of MS/MS data used a Glu-C protease specificity allowing two missed cleavages, peptide/fragment mass tolerances of 10 ppm/0.08 Da, a fixed modification of carbamidomethylation of cysteine residues, and variable modifications of acetylation at the protein N-terminus and oxidation at methionine.

5.2.6 Statistical Modeling of Digested Peptide Library

Areas of interest in the bioactivity profile were selected for subsequent data reduction and modeling. The bioactivity region for this dataset was defined as fractions 35 - 39. Using the PepSAVI-MS software package (https://cran.r-project.org/package=PepSAVIms)[15], background ions were eliminated through retention time (14 – 45 minutes), mass (200 – 1600 for *E. faecalis* digest), and charge-state (1-10, inclusive) filters to reduce the data to potential compounds of interest. For intact peptide libraries, singly charged species excluded to further select for highly positively charged bioactive peptides, however, these compounds are allowed to remain in the model for library digests. Retention time filters were selected to eliminate background ions, mass filters to select for the common mass range of bioactive peptides, and charge state filters to eliminate unwanted small molecules. Peak-picked data were binned and filtered using the previously established workflow-informed criteria [15]. Briefly, binning was performed with a 0.05 Da window of features with identical charge states and filtering required a maximum abundance inside the bioactivity area of interest (fractions 35 – 39) with <1% of that abundance outside of the chosen window. All features required minimum abundance of 100. All *m/z* species meeting these filtering criteria were modeled using the elastic net estimator with a quadratic penalty parameter specification of 0.001 to determine each species' contribution to the observed overall bioactivity profile. The resulting list contains candidate compounds ranked in order of when they

entered the model, such that the highest ranked compounds have the greatest likelihood of contributing to the bioactivity.

5.2.7 LC-MS Analysis of Intact Peptide Library

A select subset of fractions (32-42) including and surrounding the observed bioactivity region were subject to direct infusion on a Thermo Orbitrap Q Exactive HF-X for intact mass analysis. Fractions were prepared in 50% water, 50% methanol, and 0.1% formic acid with no dilution from the original library concentration, and were injected at a flow rate of 5 μL/minute. The mass spectrometer was operated at a resolving power of 120,000, positive ion mode, with 1000 – 2000 *m/z* range, and collecting 100 scans/sample. Progenesis QI for proteomics was used to generate a deisotoped peak list for intact samples, as described above.

5.2.8 Statistical Modeling of Intact Peptide Library

Exported peptide ion data for the intact library was processed as described for the digested peptide library with the following adjustments: (1) binning was performed using a mass range of 2000 – 10,000 Da to account for intact peptides and (2) a minimum intensity of 10,000,000 was required in the filtering stage to account for microscale direct infusion intensities. Exported peptide ion data contained 240 unique features, which were reduced to 177 after binning and 24 after filtering. The remaining 24 compounds entered to penalized linear regression model to determine the top 20 compounds most-likely contributing to the observed bioactivity profile.

5.2.9 Bac-21 Top-down Characterization

For top-down analysis, the most abundant Bac-21 containing *E. faecalis* pPD1 fraction was analyzed on an LTQ-Orbitrap XL platform (Thermo Fisher Scientific). The sample was subject to direct infusion utilizing a 35 micron ESI emitter (New Objective Inc.). Samples were diluted to total peptide concentration of ~ 4 μM in 80% acetonitrile, 19% water, 1% acetic acid and

injected at a flow rate of 0.5 μL/minute. The mass spectrometer was operated at a resolving power of 30,000 at 400 *m/z*, positive ion mode, with 900 – 2000 *m/z* range. The spectra were deconvoluted using the AutoXtract algorithm in Protein Deconvolution 4.0 (Thermo Fisher Scientific). CID fragmentation was performed on the +6 Bac-21 charge state (1192 *m/z*) with a collisional energy of 35 V and 600 – 2000 *m/z* range. Data analysis was accomplished using a custom informatics search engine adapted from Plymire *et al.*[21]. The candidate sequences were tested against the fragmentation data at 15 ppm mass tolerance. The searches tested fragments against every possible initial cleavage event position for the cyclic peptide tested. Outputs report the number of fragment hits, rank initial cleavage sites by a Poisson-based p-score[22], and output maps associated with each position.

5.3 Results and Discussion

5.3.1 Overview

PepSAVI-MS implements a multipronged approach for bioactive peptide discovery that utilizes selective extraction and fractionation of peptides from source material, bioactivity screening, and mass spectrometry-based peptidomics for the identification of putative bioactive peptide targets. PepSAVI-MS was originally established for constitutively-expressed peptides from botanical-sourced species and was recently validated for fungal secretome capture[23]. Now, we extend this pipeline to capture secreted peptides from bacterial sources (Figure 5.1). Additionally, we demonstrate that digestion of source libraries post-bioassay is an optional step for more facile detection of very large bioactive species (peptides/proteins) that may be difficult to detect intact depending on available LC-MS/MS hardware. Expansion of PepSAVI-MS to bacterial secretomes using digested peptide libraries highlights the versatility and wide applicability of this pipeline. Furthermore, using the secretome capture methods presented herein

in conjunction with high resolution, high mass accuracy mass spectrometry we are able to detect and characterize the active, mature Bac-21 peptide for the first time.

5.3.2 Secretome Capture and Peptide Library Generation

Dynamic microbe-microbe interactions and environmental competitive advantage are mediated by an array of secreted host-defense peptides. Although created for survival, the innate bioactivity of these peptides poses a promising source of molecules to exploit for drug discovery. As such, we have expanded PepSAVI-MS to capture potential bioactive components from the microbial secretome. Secreted peptides are collected from the cell-free supernatant using weak cation exchange (WCX) resin added directly to the media and eluted in bulk using high salt. As the majority of bioactive peptides have been found to be highly positively charged, these peptides will be retained using WCX resin. However, other resin chemistries can easily be substituted at this stage for capture of a specific class of bioactive peptides with unique chemical properties. Concentrated elutions are then combined and crudely fractionated using SCX for peptide library creation.

5.3.3 Bioactivity Screening

PepSAVI-MS is amenable to a variety of bioassay formats and can be modified to accommodate any target pathogen. High-throughput microtiter-based assays and agar diffusion-based assays have been previously demonstrated with PepSAVI-MS[15, 23]. While many bacterial species bode well in high throughput 96-well assays (as presented in the original implementation of the PepSAVI-MS pipeline), this format is not amenable to some bacterial species that fail to grow to high density (e.g. < 0.5 OD) in low volume formats. This was the case with enterococcal cultures in Mueller Hinton Broth (MHB) and as such, agar-based diffusion assays (originally developed for fungal cultures) were used to examine activity of the *E. faecalis* peptide library

against its respective susceptible culture, *E. faecium*. The *E. faecalis* pPD1 peptide library shows strong activity against *E. faecium*, with measurable zones of inhibition present in fractions 35 – 39 (Figure 5.2). The bioactivity region is a discrete and well-defined profile spanning five sequential fractions with varying zones of inhibition. This representative profile is ideal for PepSAVI-MS as the bioactive peptide responsible for the activity will have five paired bioactivity and LC-MS abundance data points to facilitate statistics and modeling.

5.3.4 MS Profiling of *E. faecalis* pPD1 Peptide Library

The peptide library is subjected to LC-MS/MS analysis to obtain accurate mass and peptide abundance data for all constituents. PepSAVI-MS was originally demonstrated with top-down, intact peptidome libraries. However, an optional digestion prior to LC-MS/MS analysis may facilitate detection of target peptides in a manner similar to the bottom-up Activity-Correlated Quantitative Proteomics Platform (ACPP) used to identify enzymes with targeted reaction specificity[24]. This approach is particularly useful for larger peptides with low charge density or poor ionization efficiency, or for investigations implementing resolving power-limited mass spectrometry hardware. In such cases, database searching against the same or highly related microbial species can be used to determine the peptide constituents within each fraction and the proteolytic peptide(s) can be used for quantification and modeling via PepSAVI-MS. Proteolytic peptides highly ranked as putatively bioactive can be used as a proxy for the intact peptide present in each fraction. This digestion step is demonstrated herein for the *E. faecalis* peptide library, and is validated through comparison to the original top-down intact analysis approach. Positively charged residues (i.e. Arg, Lys) are frequently occurring in most bioactive peptides and would be cleaved by trypsin into many short peptides that elude LC-MS detection. Thus, endoproteinase Glu-C, which cleaves c-terminal of glutamic acid residues, was used for

proteolytic processing in order to obtain longer peptides with a higher probability to achieve increased protein coverage.

Peptide ion data reveals 10,132 features detected across Glu-C digested fractions 30-45 for the *E. faecalis* pPD1 library. This peptide ion data is used for both unbiased data reduction and statistical modeling with unassigned *m/z* ratios and database searching using Mascot for identification. For unbiased data reduction and statistical modeling, peptide ion data is first reduced to features most-likely contributing to the bioactivity in each region through mass, retention time, and elution profile mapping. When applied to the *E. faecalis* pPD1 data set, this reduction identified 3,727 unique *m/z* ratios that were further filtered to 23 compounds with abundance profiles spanning the bioactivity region. Statistical modeling using elastic net penalized linear regression of this filtered data set reveals a candidate list of compounds likely responsible for the bioactivity region. After modeling, a list of the top 20 m/z's likely contributing to the observed bioactivity is generated (Figure 5.3). Then, comparison of the top 20 contributing *m/z* ratios with the Mascot output for the entire dataset facilitates the assignment of top contributors. In this case, the 2nd, 4th and 5th top contributors (*m/z* ratios 1555.86, 540.32 and 360.55) correspond to the +2 and +3 charge states of two Bac-21 Glu-C peptides: FGIPAAVAGTVLNVVE (+2) and SIKAYLKKE (+2, +3) (Figures 5.3 and 5.4). Glu-C digestion of the *E. faecalis* pPD1 library yielded three of the four possible Bac-21 peptides, providing 58% sequence coverage of the mature peptide. While three peptides were detected using Mascot, only two (FGIPAAVAGTVLNVVE and SIKAYLKKE) ranked in the top 20 compound list after statistical modeling. It is likely that low abundance and incomplete digestion were contributing factors to this result. Enzymatic digestion using proteases with higher

digestion efficiency may increase the number of identified and ranked proteolytic peptides, increasing confidence in bioactive peptide identification in future experiments.

Analysis of intact *E. faecalis* pPD1 peptide library reveals 240 unique features across fractions 32-42. Binning and filtering parameters were adjusted for the increased size of intact peptides, but was modeled against the same bioactivity data as the Glu-C digested peptide library. The resulting top 20 list after binning, filtering, and modeling featured all three of the detected Bac-21 charge states, ranking 2nd (+6), 14th (+4) and 17th (+5).

5.3.5 Characterization of Intact Bac-21

While evidence of *E. faecalis* Bac-21 has been demonstrated through bioactivity[14], Bac-21 has never been molecularly characterized. Hence, the most abundant Bac-21 containing library fraction was characterized pre-digestion (i.e. intact) using direct infusion, high resolution mass spectrometry in combination with collision induced dissociation for increased sensitivity and accurate mass detection. As intact Bac-21 is a cyclic peptide, two fragmentation events must occur to produce fragment ions via MS/MS. The initial fragmentation is a ring opening event that can occur at any amino acid residue but will produce the same precursor mass. The second fragmentation event produces the classical b- and y- ion series typically associated with CID. While the initial fragmentation event produces identical linear masses independent of the initial cleavage point, the b- and y-ions generated from each of these linear sequences will be unique to the initial cleavage point. As such, fragmentation of cyclic peptides exponentially increases the number of possible fragment ions produced from a single peptide. While providing a wealth of information to use for peptide sequencing, interpretation of these MS/MS spectra is often complex and confident fragment assignment is often ambiguous in many cases. Automatic software programs are crucial to facilitate this process, and an in-house custom informatics

search engine[21] was used in this case. As such, the accurate intact mass of Bac-21 was determined to be 7145.071-*0* Da (7145.072-*0* Da theoretical; 0.3 ppm mass accuracy) with 70% sequence coverage obtained via CID fragmentation (Figure 5.5). These results indicate that Bac-21 is identical in both primary protein sequence and post-translational processing (head-to-tail cyclization) to the known bacteriocin AS-48[11, 13]. Unsurprisingly, the most frequent initial fragmentation event occurred at the only proline residue (denoted as the 8th amino acid) in Bac-21. The next, most frequent initial fragmentation events occurred at 11th, 12th, and 36th amino acid residues of V, A, and G, respectively, and likely indicate structural weak points in the folded peptide.

5.4 Conclusion

PepSAVI-MS is a highly versatile and easily adaptable pipeline for natural product bioactive peptide discovery. Successful application of PepSAVI-MS for bacterial secretome capture demonstrates rapid and accurate identification of bioactive components from complex natural product secretomes and allows for focused downstream characterization and validation experiments on only the most promising species. Expansion of PepSAVI-MS to this historically rich source of bioactive molecules greatly enhances its capabilities and potential for novel bioactive peptide discovery. Furthermore, minor modifications to the original pipeline, including the use of agar diffusion assays for recalcitrant bacterial species and digestion of peptide libraries for sequenced organisms, facilitates exploration of bacterial secretomes and expands the applicability of PepSAVI-MS to data acquired with lower resolution mass spectrometry hardware. Successful application of PepSAVI-MS to microbial secretomes, as demonstrated with *E. faecalis*, opens the door to investigating microbial species with a new lens and stands to have positive implications in both human health and agriculture.

5.5 Figures

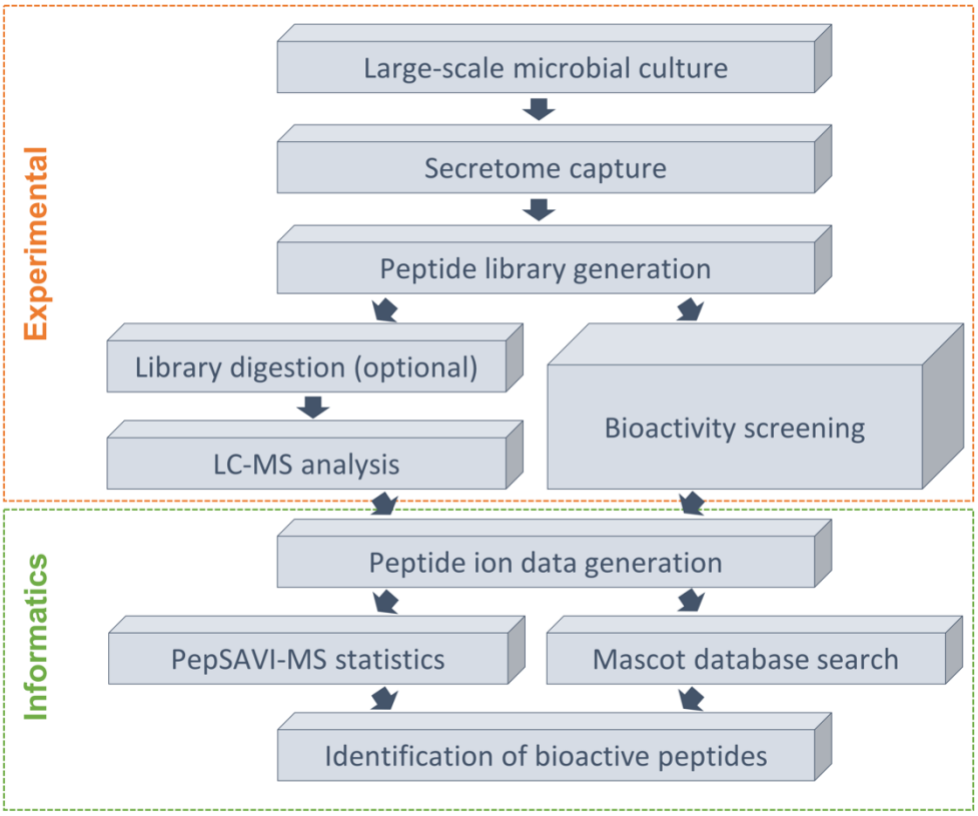


Figure 5.1 Overall workflow for PepSAVI-MS application to bacterial secretomes using digested peptide libraries. The developed workflow is amenable to large-scale microbial culture growth to produce secreted peptides in sufficient quantities for downstream analysis. Secreted peptides are captured with cation exchange resin and crudely fractionated for creation of each peptide library. Peptide libraries are subject to LC-MS analysis and bioactivity screening, and generated data sets are combined and informatically processed to identify bioactive peptides.

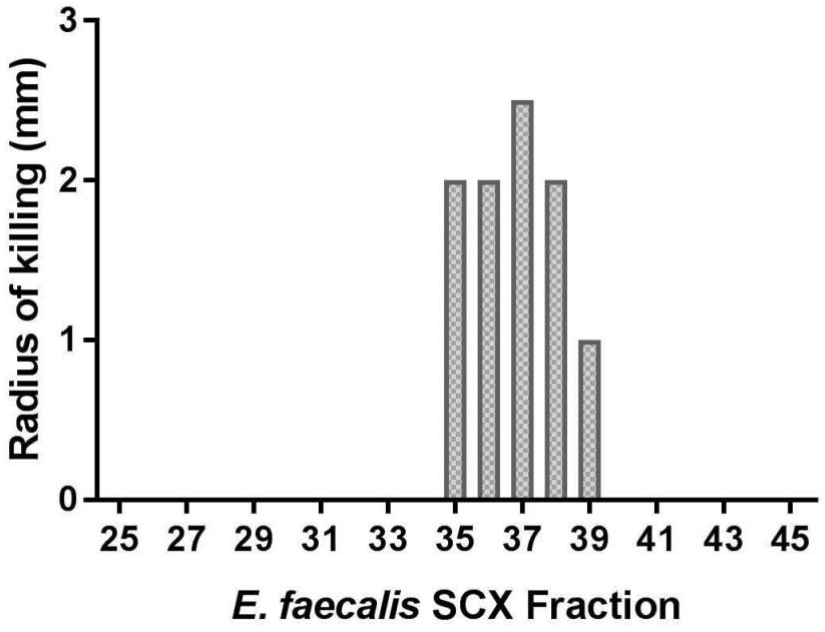


Figure 5.2 *E. faecalis* pPD1 SCX fraction library bioactivity against *E. faecium* JL282. Assays were performed in duplicate and the average radius of inhibition is plotted for each fraction.

Rank	(*m/z*)	Charge	Mass	Bac-21 Peptide
1	661.643	1	660.635	
2	778.940	2	1555.864	FGIPAAVAGTVLNVVE (+2)
3	773.443	1	772.435	
4	540.322	2	1078.628	SIKAYLKKE (+2)
5	360.554	3	1078.638	SIKAYLKKE(+3)
6	1079.649	1	1078.641	
7	379.561	3	1135.661	
8	398.552	3	1192.632	
9	1193.693	1	1192.685	
10	597.348	2	1192.681	
11	493.548	2	985.081	
12	378.360	3	1132.056	
13	554.324	2	1106.633	
14	375.218	3	1122.630	
15	365.886	3	1094.633	
16	625.860	2	1249.704	
17	359.412	3	1075.211	
18	1136.671	1	1135.663	
19	364.553	3	1090.637	
20	568.838	2	1135.659	

Figure 5.3 PepSAVI-MS statistical modeling summary, including the list of the top 20 compounds, denoted by rank, m/z ratio and charge state, to exit the PepSAVI-MS statistical model for the digested peptide library. The first compound can be considered the most likely compound contributing to the observed bioactivity, with the likelihood decreasing down the rank list. Ranked peptides identified as Bac-21 are noted.

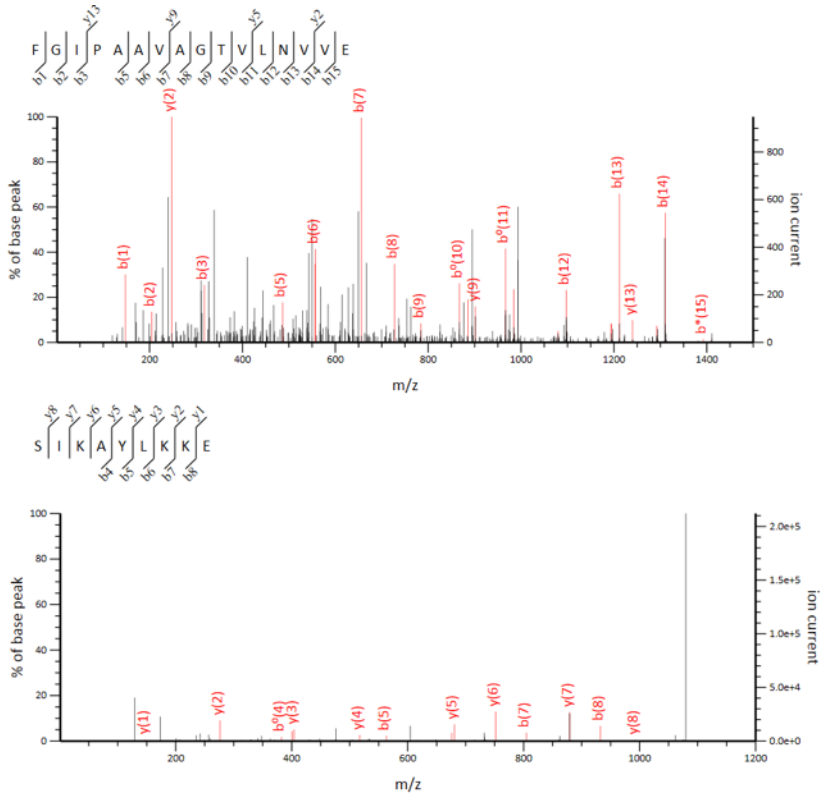


Figure 5.4 Representative MS/MS spectra generated from CID fragmentation of Bac-21 Glu-C peptides on a TripleTOF 5600 Q-TOF platform. Peptide fragment maps used for Mascot identification of each statistically ranked Bac-21 peptide are shown.

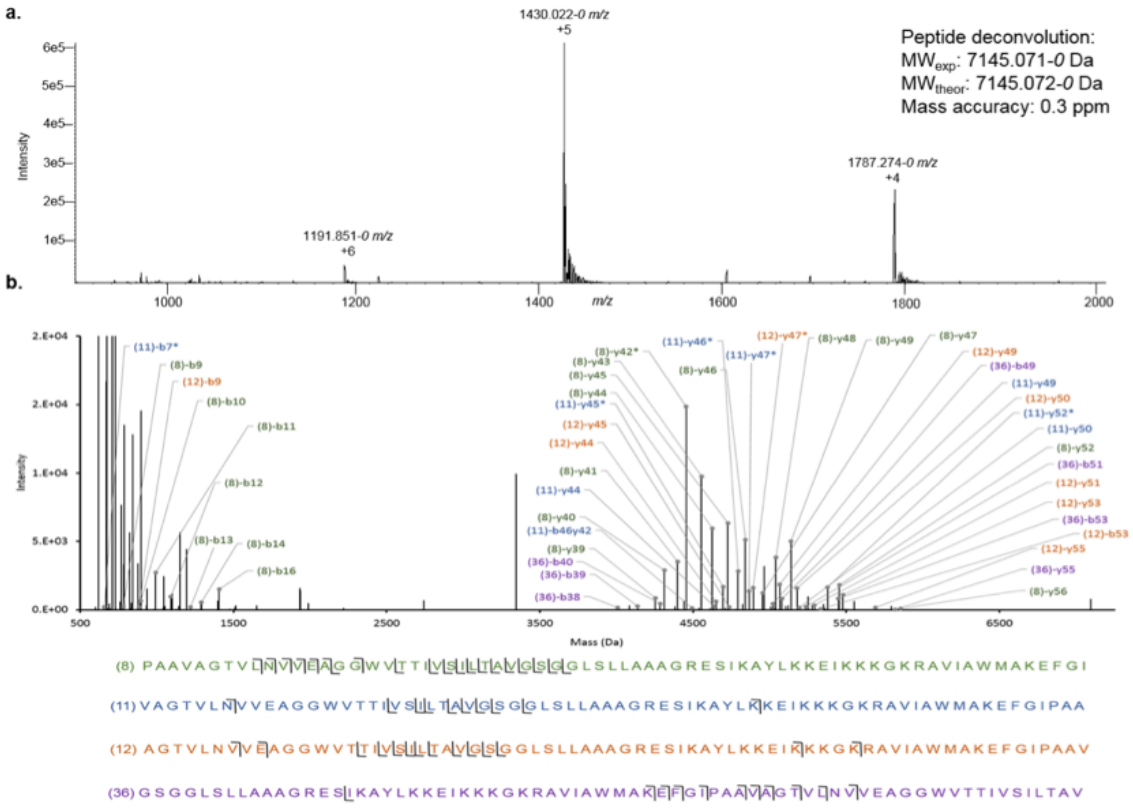


Figure 5.5 (a) Mass spectrum of intact Bac-21 (LTQ-Orbitrap XL). (b) Deconvoluted and annotated CID spectrum of 1192 *m/z* (Bac-21 +6 charge state). Fragmentation coverage was determined using the four most prevalent linear sequences, denoted by the number of the starting amino acid from the initial cyclic sequence. B- and y-ions are labeled using traditional nomenclature with the number of the starting amino acid residue in parenthesis. Full linear sequences and ion coverage are provided below for clarity. Ions denoted with * indicate that another fragment identity is possible for that detected mass.

REFERENCES

1. Huycke, M. M.; Sahm, D. F.; Gilmore, M. S., Multiple-drug resistant enterococci: the nature of the problem and an agenda for the future. *Emerg. Infect. Dis.* **1998,** *4* (2), 239-249.

2. Kristich, C. J.; Rice, L. B.; Arias, C. A., Enterococcal Infection-Treatment and Antibiotic Resistance. In *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*, Gilmore, M. S.; Clewell, D. B.; Ike, Y.; Shankar, N., Eds. Massachusetts Eye and Ear Infirmary: Boston, **2014**.

3. Miller, W. R.; Munita, J. M.; Arias, C. A., Mechanisms of antibiotic resistance in enterococci. *Expert Rev. Anti Infect. Ther.* **2014,** *12* (10), 1221-1236.

4. Olawale, K. O.; Fadiora, S. O.; Taiwo, S. S., Prevalence of Hospital-Acquired Enterococci Infections in Two Primary-Care Hospitals in Osogbo, Southwestern Nigeria. *African Journal of Infectious Diseases* **2011,** *5* (2), 40-46.

5. Corr, S. C.; Li, Y.; Riedel, C. U.; O'Toole, P. W.; Hill, C.; Gahan, C. G., Bacteriocin production as a mechanism for the antiinfective activity of Lactobacillus salivarius UCC118. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104* (18), 7617-7621.

6. Perez, R. H.; Zendo, T.; Sonomoto, K., Novel bacteriocins from lactic acid bacteria (LAB): various structures and applications. *Microb Cell Fact* **2014,** *13 Suppl 1*, S3.

7. Grande Burgos, M. J.; Pulido, R. P.; Del Carmen Lopez Aguayo, M.; Galvez, A.; Lucas, R., The Cyclic Antibacterial Peptide Enterocin AS-48: Isolation, Mode of Action, and Possible Food Applications. *Int. J. Mol. Sci.* **2014,** *15* (12), 22706-22727.

8. Cruz, V. L.; Ramos, J.; Melo, M. N.; Martinez-Salazar, J., Bacteriocin AS-48 binding to model membranes and pore formation as revealed by coarse-grained simulations. *Biochim. Biophys. Acta* **2013,** *1828* (11), 2524-2531.

9. Dobson, A.; Cotter, P. D.; Ross, R. P.; Hill, C., Bacteriocin production: a probiotic trait? *Appl. Environ. Microbiol.* **2012,** *78* (1), 1-6.

10. Hibbing, M. E.; Fuqua, C.; Parsek, M. R.; Peterson, S. B., Bacterial competition: surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.* **2010,** *8* (1), 15-25.

11. Tomita, H.; Fujimoto, S.; Tanimoto, K.; Ike, Y., Cloning and genetic and sequence analyses of the bacteriocin 21 determinant encoded on the Enterococcus faecalis pheromone-responsive conjugative plasmid pPD1. *J. Bacteriol.* **1997,** *179* (24), 7843-7855.

12. Wardal, E.; Sadowy, E.; Hryniewicz, W., Complex nature of enterococcal pheromone-responsive plasmids. *Pol. J. Microbiol.* **2010,** *59* (2), 79-87.

13. Maqueda, M.; Galvez, A.; Bueno, M. M.; Sanchez-Barrena, M. J.; Gonzalez, C.; Albert, A.; Rico, M.; Valdivia, E., Peptide AS-48: prototype of a new class of cyclic bacteriocins. *Curr Protein Pept Sci* **2004,** *5* (5), 399-416.

14. Kommineni, S.; Bretl, D. J.; Lam, V.; Chakraborty, R.; Hayward, M.; Simpson, P.; Cao, Y.; Bousounis, P.; Kristich, C. J.; Salzman, N. H., Bacteriocin production augments niche competition by enterococci in the mammalian gastrointestinal tract. *Nature* **2015,** *526* (7575), 719-722.

15. Kirkpatrick, C. L.; Broberg, C. A.; McCool, E. N.; Lee, W. J.; Chao, A.; McConnell, E. W.; Pritchard, D. A.; Hebert, M.; Fleeman, R.; Adams, J.; Jamil, A.; Madera, L.; Stromstedt, A. A.; Goransson, U.; Liu, Y.; Hoskin, D. W.; Shaw, L. N.; Hicks, L. M., The "PepSAVI-MS" Pipeline for Natural Product Bioactive Peptide Discovery. *Anal. Chem.* **2017,** *89* (2), 1194-1201.

16. Riley, M. A.; Wertz, J. E., Bacteriocins: evolution, ecology, and application. *Annu. Rev. Microbiol.* **2002,** *56*, 117-137.

17. Golneshin, A.; Adetutu, E.; Ball, A. S.; May, B. K.; Van, T. T.; Smith, A. T., Complete Genome Sequence of Lactobacillus plantarum Strain B21, a Bacteriocin-Producing Strain Isolated from Vietnamese Fermented Sausage Nem Chua. *Genome Announc* **2015,** *3* (2), e00055-00015.

18. Kristich, C. J.; Little, J. L., Mutations in the beta subunit of RNA polymerase alter intrinsic cephalosporin resistance in Enterococci. *Antimicrob. Agents Chemother.* **2012,** *56* (4), 2022-2027.

19. Yagi, Y.; Kessler, R. E.; Shaw, J. H.; Lopatin, D. E.; An, F.; Clewell, D. B., Plasmid content of Streptococcus faecalis strain 39-5 and identification of a pheromone (cPD1)-induced surface antigen. *J. Gen. Microbiol.* **1983,** *129* (4), 1207-1215.

20. Vizcaino, J. A.; Csordas, A.; del-Toro, N.; Dianes, J. A.; Griss, J.; Lavidas, I.; Mayer, G.; Perez-Riverol, Y.; Reisinger, F.; Ternent, T.; Xu, Q. W.; Wang, R.; Hermjakob, H., 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **2016,** *44* (D1), D447-456.

21. Plymire, D. A.; Wing, C. E.; Robinson, D. E.; Patrie, S. M., Continuous Elution Proteoform Identification of Myelin Basic Protein by Superficially Porous Reversed-Phase Liquid Chromatography and Fourier Transform Mass Spectrometry. *Anal. Chem.* **2017,** *89* (22), 12030-12038.

22. Meng, F.; Cargile, B. J.; Miller, L. M.; Forbes, A. J.; Johnson, J. R.; Kelleher, N. L., Informatics and multiplexing of intact protein identification in bacteria and the archaea. *Nat. Biotechnol.* **2001,** *19* (10), 952-957.

23. Kirkpatrick, C. L.; Parsley, N. C.; Bartges, T. E.; Cooke, M. E.; Evans, W. S.; Heil, L. R.; Smith, T. J.; Hicks, L. M., Fungal Secretome Analysis via PepSAVI-MS: Identification of the Bioactive Peptide KP4 from Ustilago maydis. *J. Am. Soc. Mass. Spectrom.* **2018**.

24. Ma, H.; Delafield, D. G.; Wang, Z.; You, J.; Wu, S., Finding Biomass Degrading Enzymes Through an Activity-Correlated Quantitative Proteomics Platform (ACPP). *J. Am. Soc. Mass. Spectrom.* **2017,** *28* (4), 655-663.

CHAPTER 6: Discovery of a Novel Antibacterial Peptide from *Amaranthus tricolor* using PepSAVI-MS

*Manuscript to be submitted for publication. Authors: Kirkpatrick, C.L., Bartges, T.E., Heil, L.R., and Hicks, L.M.

6.1 Introduction

Plants of the Amaranthus spp. have long been known as natural remedies for chronic illnesses. Amaranthus extracts have been used across Southern and Eastern Asia to treat conditions including urinary infections, diarrhea, pain, respiratory disorders, and diabetes[1-3]. Now, with the growing antimicrobial resistance concern, these ancient remedies are receiving increased attention a viable source of therapeutic compounds. Due to this reinvigorated effort, recent studies have revealed anticancer, antiviral, hepatoprotective, neuroprotective, cardioprotective, and antidiabetic activities of Amaranthus[1].

Recent research exploring Amaranthus has utilized virtually all parts of the plant under different extraction conditions and testing for different antioxidant and pharmacological activities[1]. However, the majority of these studies implemented organic extractions and largely focused on small molecule constituents. To our knowledge, most of the reported proteinaceous activity of Amaranthus is from the seeds of various species[4-8]. Activity detected from the seeds in these studies includes anticancer and antihypertenstion bioactivity. While exciting, we posit that peptidyl activity across Amaranthus spp. remains underexplored, especially within aerial tissues. As such, we apply our bioactive peptide discovery pipeline, PepSAVI-MS, to *Amaranthus tricolor* to probe proteinaceous antimicrobial activity.

Amaranthus tricolor, commonly known as edible amaranth, Joseph's coat, or red spinach, is of particular interest due to its history in traditional medicine. The whole plant has been consumed for centuries in Indian culture to promote general overall health, and the roots have been used to control bleeding[9]. Recent research has revealed the leaves of *A. tricolor* to be anti-nociceptive, antidiabetic, hepatoprotective, gastroprotective, neuroprotective and anti-inflammatory[10-16]. *A. tricolor* leaf powder has even been shown to prevent post-menopause complications, highlighting antioxidant activity of these tissues[17]. While extremely diverse activity has been identified, antimicrobial action of this particular species is largely unexplored. Furthermore, the activity of *A. tricolor* extracts has been attributed to the extract as a whole (i.e. specific active compound not identified) or attributed to small molecules. As such, innovative tools that specifically target bioactive peptide discovery are required to probe this unique subset of compounds.

The PepSAVI-MS pipeline has been developed to expedite the search for natural product bioactive peptides and has been validated for use in plants through the identification of the bioactive peptide cyO2 from *Viola odorata*[18]. Herein, we explore antimicrobial activity from the leaves of *A. tricolor* using PepSAVI-MS. Bioactivity results indicate strong activity against the gram-negative bacterium *Escherichia coli*. This activity is retained through both SCX and reversed phase fractionation, indicating high stability of the target peptide. Using the PepSAVIms statistical analysis package, the top 20 candidates most-likely contributing to the observed bioactivity have been identified. Further work is being performed to identify the specific bioactive component, and to isolate or synthesize the active compound. The purified bioactive peptide will undergo activity validation against *E. coli* and will be screened for activity

against a full panel of gram-negative and gram-positive bacteria to assess full bioactivity potential.

6.2 Materials and Methods

6.2.1 Plant Growth and Harvest

Amaranthus tricolor seeds purchased from Strictly Medicinal Seeds (formerly Horizon Herbs; Williams, OR) were grown to mature rosette stage (~8 weeks) (Figure 6.1). Approximately 100 seeds were planted in nutrient-rich soil and grown under controlled temperature (63.5 - 68.5°F) and light cycle (14-hr light) conditions. Due to the high germination rate, seedlings were thinned after two weeks to allow ample space and nutrients to for growth. Following an additional two weeks, remaining healthy seedlings were transplanted into individual pots. Aerial tissue was taken from ~20 healthy, mature plants (equivalent to ~140 g of powdered material), harvested, immediately flash frozen with liquid nitrogen, and stored at -80°C until subsequent extraction.

6.2.2 Peptide Library Creation

Peptide Extraction. Frozen tissue (100 g) was ground under liquid nitrogen using a mortar and pestle and aqueously extracted in 300 mL of 10% acetic acid with protease inhibitors (Roche, 6 tablets) with stirring for 4 hours at 4°C. Insoluble material was pelleted by centrifugation and filtered using a 0.45 μm Stericup filtration (Millipore). Protein concentration filters (30 kDa, Millipore) were used to remove high molecular weight species followed by dialysis (0.1 – 1 kDa cutoff; SpectrumLabs) into 5 mM ammonium formate pH 2.7 to eliminate small molecules. The sample was concentrated via vacuum centrifugation to 4 mL to generate the final crude extract.

HPLC Fractionation. Four crude extracts were combined and further concentrated to 420 μL to create a 4x concentrated crude extract. The concentrated extract was subjected to a 47-minute SCX method using a PolySulfoethylA column (100 x 4.6 mm, 3 μm particles, PolyLC). A salt gradient was employed using a linear ramp from 5 mM ammonium formate, 20% acetonitrile, pH 2.7 to 500 mM ammonium formate, 20% acetonitrile, pH 3.0. Fractions were collected across the 47-minute run in one-minute intervals and desalted with three washes of 1.3 mL water using a vacuum concentrator. Peptide libraries were stored in water (250 μL/fraction) at 4°C until bioactivity assay.

6.2.3 Bioactivity Screening

Escherichia coli (25922) was obtained through ATCC and the ESKAPE pathogen strains are clinical isolates that belong to a collection acquired by the Shaw Lab [*Staphylococcus aureus* (635), *Klebsiella pneumoniae* (1433), *Acinetobacter baumannii* (1403)][19]. Bacterial starter cultures were grown from an isolated colony on a freshly streaked agar plate in 5 mL of Mueller Hinton Broth (MHB). Overnight cultures were diluted to 0.1 OD in the morning and grown for one additional hour. After, a master mix was prepared by combining 4 mL of 1x MHB, 2 mL of 2x MHB, and 2 mL of bacterial culture, such that the plating OD was ~0.1. All assays were performed in 96-well microtiter plates by adding 10 μL of peptide fraction to 40 μL master mix. Positive (ampicillin, 100 μg/mL) and negative (water) controls were added to each plate. The prepared plates were incubated (37°C, 275 rpm shaking) for 4 hours depending before the addition of 1 μL resazurin (50 mM) to each well. After one additional hour of shaking and incubation, a fluorescence read of 544 nm (ex) and 590 nm (em) was collected to measure relative fluorescence units for each sample, and percent activity of each well was calculated.

6.2.4 Reversed Phase Separation

As the initial SCX peptide library displayed activity in an unusually large number of fractions, a second round of fractionation was used to further separate and probe individual regions of activity. Active SCX fractions 27, 28, and 29 were combined (~150 µL/fraction) for further reversed phase fractionation. 420 µL of combined sample was subject to reversed phase fraction on a Shimadzu HPLC using a Phenomenex Jupiter C18 column (150 x 4.6 mm, 5 µm, 300 Å). Aqueous (95% water/5% acetonitrile/0.1% TFA) and organic (100% acetonitrile/0.1% TFA) solutions were used for mobile phase A and B, respectively. A two-step organic gradient was applied, and 1-minute fractions were collected across the gradient. Fractions were dried down and resuspended in 25 µL of water. Reversed phase (RP) fractions were subject to bioactivity screening as described above, with the exception of being in conducted in duplicate instead of triplicate to conserve material.

6.2.5 LC-MS/MS Analysis

A. tricolor SCX and RP peptide libraries (fractions 11-40 and 11-35, respectively) were analyzed on a nano-LC-ESI-MS/MS platform composed of a Waters nanoAcquity UPLC coupled to an AB Sciex TripleTOF 5600 QTOF mass spectrometer. Peptide fractions were diluted to ~0.2 µg/µL and acidified to 0.1 % formic acid. Five microliters of each sample were injected onto a trap column (NanoAcquity UPLC 2G-W/M Trap 5 µm Symmetry C18, 180 µm x 20 mm: Waters) before transfer to the analytical C18 column (10k PSI, 100 Å, 1.8 µm, 75 µm x 250 mm: Waters). Peptide separation was carried out at a flow rate of 0.3 µL/minute using a linear ramp of 5 – 50 % mobile phase B (mobile phase A, 0.1% formic acid; mobile phase B, 0.1% formic acid in acetonitrile) over 30 minutes. The MS was operated in positive ion, high sensitivity mode with the MS survey spectrum using a mass range of 350-1600 *m/z* in 250 ms

and information dependent acquisition of MS/MS data, 87 ms per scan. For IDA MS/MS experiments, the first 20 features above 150 counts threshold and having a charge state of +2 to +5 were fragmented using rolling collision energy ±5%. Each MS/MS experiment put the precursor *m/z* on an 8-second dynamic exclusion list. Auto calibration was performed every eight samples (8 h) to assure high mass accuracy in both MS and MS/MS acquisition. Deisotoped peak lists for each library were generated using Progenesis QI for Proteomics software (Nonlinear Dynamics, v.2.0). Automatic processing settings were used to align and peak pick ions across all runs, using fractions 28 and 17 as the reference runs for the SCX and RP libraries, respectively. Additionally, a minimum peak with of 0.1 minutes was applied. Identified features were quantified using AUC integration of survey scan data based on the summed intensity of each deisotoped feature. Data was exported as "peptide ion data" with the default parameters from Progenesis at the "Identify Peptides" stage in the software.

6.2.6 Statistical Modeling of Reversed Phase Fractions

The bioactivity region for *A. tricolor* vs. *E. coli* was defined based on the observed bioactivity profile as fractions 16 - 19. Using the PepSAVI-MS software package (https://cran.r-project.org/package=PepSAVIms)[18], background ions were eliminated through retention time (14 – 45 minutes), mass (1000 – 10000 Da), and charge-state (2-10, inclusive) filters to reduce the data to potential compounds of interest. Peak-picked data were binned and filtered using the previously established workflow-informed criteria[18]. Briefly, binning was performed with a 0.05 Da window of features with identical charge states and filtering required a maximum abundance inside the extended bioactivity area of interest (16 – 19) with <10% of that abundance outside of the chosen window. All features required a minimum abundance of 100. All *m/z* species meeting these filtering criteria were modeled using the elastic net estimator with a quadratic penalty

parameter specification of 0.001 to determine each species' contribution to the observed overall bioactivity profile. The resulting list contains candidate compounds ranked in order of when they entered the model, such that the highest ranked compounds have the greatest likelihood to be contributing to the bioactivity.

6.2.7 Peptide Characterization

The most promising peptide target from the top 20 modeled species is undergoing subsequent characterization for primary sequence determination. The *A. tricolor* SCX fraction with the highest abundance of target peptide was split into three samples that were reduced using 10 mM dithiothreitol (30 minutes, 45°C, 850 rpm) and alkylated with 100 mM iodoacetamide (15 minutes, 25°C, 850 rpm). One aliquot was quenched and frozen at this point, and the other two were subject to overnight digestion with either trypsin or Glu-C (37°C, 850 rpm). *A. tricolor* digested fractions were cleaned up with Pierce C18 zip tips (Thermo Fisher Scientific) before subsequent LC-MS analysis. The reduced and alkylated sample was used to identify predicted mass shifts associated with disulfide bonds. The digested samples were searched using Mascot (v.2.5.0; Matrix Science, http://www.matrixscience.com/) for protein identification. Database searching was performed using the *A. tricolor* database generated from transcriptomic data (21,758 entries; accessed November, 2017). Searches of MS/MS data used either trypsin or Glu-C protease specificity allowing two missed cleavages, peptide/fragment mass tolerances of 10 ppm/0.08 Da, a fixed modification of carbamidomethylation of cysteine residues, and variable modifications of acetylation at the protein N-terminus and oxidation at methionine.

6.3 Results and Discussion

6.3.1 Overview

Amaranthus tricolor has long been known for its role as a traditional medicinal species with a wealth of healing capabilities. Broad spectrum activity of this species has largely been attributed to small molecules ranging from phenolic compounds to betalains[4-8], however activity due to peptidyl components remains unexplored. Here, we use PepSAVI-MS to probe the bioactive peptide pool from *A. tricolor* and identify novel lead compounds for identification and activity validation.

6.3.2 Plant Growth and Peptide Extraction

A. tricolor seeds were grown under standard greenhouse conditions for 8 weeks until ample aerial tissue was available. Healthy leaves were harvested from the plants, packaged in aluminum foil packets, and flash frozen under liquid nitrogen. Peptide extraction was performed using 100 g of ground aerial tissue extracted into an aqueous buffer, as described by Kirkpatrick *et al*[18]. Size exclusion filters were used to select for peptide-sized compounds (1-10 kDa), and sample was concentrated for creation of the *A. tricolor* crude extract.

6.3.3 Peptide Library Creation

The *A. tricolor* crude extract was subject to strong cation exchange fractionation, where positively charged bioactive peptides retain on the column until elution with a high salt gradient. As shown in Figure 6.2a, the 280 nm SCX trace for *A. tricolor* crude extract contains high levels of unretained compounds eluting before the salt gradient is applied. These compounds are likely small, negatively charged or neutral molecules common in plant extracts. As activity of this subset of compounds has been thoroughly studied and reported in literature, the focus of this study is on later fractions (11-40) likely to contain highly basic peptides. One-minute fractions

are collected across the optimized 47-minute run, and once desalted comprise the SCX peptide library for downstream analysis.

6.3.4 Bioactivity Screening: SCX Peptide Library

A. tricolor SCX peptide library was subject to preliminary bioactivity screening versus *E. coli*. Each peptide fraction was incubated with actively growing *E. coli* cells, in triplicate, and after 4 hours of incubation the number of remaining cells in each well was quantified using the cell viability indicator, resazurin. As seen in Figure 6.2b, activity was detected across many fractions in the peptide library. This result is unusual for plant extracts and suggests that the fractions may not have been desalted thoroughly. However, clear regions of activity are still apparent, across fractions 18-24, 25-30, and 31-36. This assay has been repeated for clarity/bioactivity confirmation (data not shown), and the region with the strongest activity (Figure 6.2b, fractions 27-29, highlighted in blue) was selected for downstream analysis.

6.3.5 Reversed Phase Fractionation of Active SCX Fractions

As the initial SCX peptide library displayed activity in an unusually large number of fractions, a second round of fractionation was used to further separate and probe individual regions of activity. As such, SCX fractions with the highest activity levels (27-29) were recombined and subject to reversed phase separation and fractionation. A 42-minute optimized run with a two-step organic gradient to elute retained compounds was implemented. The resulting 220 nm trace is shown in Figure 6.3a. A small, unretained peak from 2-3 minutes is still evident; however, the majority of compounds in the sample are being retained on the reversed phase column. A strikingly wide peak is apparent from 16-19 minutes, but all other peaks display tight, sharp peak shapes characteristic of reversed phase separation. For creation of the *A.*

tricolor RP peptide library, 1-minute fractions were collected across the run, dried down, and resuspended in 25 μL of LC-MS grade water.

6.3.6 Bioactivity Screening: RP Peptide Library

Reversed phase library fractions were subject to bioactivity screening vs. *E. coli*, as described above except in duplicate due to limited sample volume. As seen in Figure 6.3b, a single, strong bioactivity region is detected across fractions 16-19 (highlighted in blue), with maximum activity occurring in fraction 17 and decreasing on either side. The additional active fraction to the right of the maximum suggests peak tailing, which aligns with the trend in UV data across that time frame. Reversed phase separation has the potential to denature peptides through elimination of secondary and tertiary structure often crucial to activity. However, conservation of activity after reverse phase fractionation indicates stability of the target peptide and is an exciting feature of candidates for drug development.

6.3.7 LC-MS and Statistical Analysis of Peptide Libraries

Both SCX and RP *A. tricolor* peptide libraries (fractions 11-40 and 11-35, respectively) were subjected to LC-MS/MS analysis. Peptide ion data reveals 11,558 and 8,001 features in the SCX and RP data sets, respectively. As the reversed phase peptide library generated a more defined bioactivity profile, PepSAVIms statistical analysis was performed on this data set. Binning of mass spectrometry features to reduce compound redundancy condensed this data set to 4,070 features, and further filtering for compounds with their maximum abundance within the bioactivity region revealed 119 features possibly contributing to the bioactivity. Penalized linear regression of these remaining features divulges the top 20 compounds most-likely contributing to the bioactivity (Figure 6.4). Manual interrogation of the top 20 compounds reveals promise in the

top candidate, 568.95 *m/z*. This feature corresponds to a 1703.82 Da peptide that was detected in three charge states and elutes around 17 minutes in the reversed phase nanoLC gradient.

6.4 Peptide Characterization

Experiments to sequence target peptides in the top 20 compound list are underway. The primary target of interest, a 1703.82 Da peptide, was selected for tandem MS in the DDA workflow, but displays minimal fragmentation. Additional measures for characterization including targeted fragmentation with increased collision energy on the TripleTOF, and complementary HCD fragmentation using the Orbitrap HF-X are underway.

Additionally, although *A. tricolor* has not been genome sequenced transcriptomic data is available as part of the oneKP initiative to generate large-scale gene sequencing data for 1000 plant species[20-23]. This data has been used for in-house translation of a protein database that can facilitate a traditional bottom-up proteomics approach. As such, a reduction, alkylation, digestion (trypsin and Glu-C) workflow is being used to achieve high-quality MS/MS data to be searched against the *A. tricolor* protein database in Mascot. Initial reduction and alkylation mass shifts indicate the presence of one disulfide bond in the target peptide. No additional mass shifts of the target peptide are detected after subsequent Glu-C digestion, indicating the absence of glutamic acid and aspartic acid residues. Tryptic digestion eliminates the reduced peptide therefore indicating the presence of multiple basic residues within the target peptide.

6.5 Bioactivity Characterization

To further probe the activity of the *A. tricolor* peptide library, SCX fractions were screened against three of the ESKAPE pathogens: *Staphylococcus aureus, Klebsiella pneumoniae,* and *Acinetobacter baumannii*. There appears to be low levels of activity in *A. tricolor* fractions 27-29

vs. *A. baumannii* (Figure 6.5) which may be similar to the activity seen vs. *E. coli*. As such, this assay is being repeated with a higher concentration of peptide library for bioactivity assessment. Additionally, more material is being prepared to screen against the remaining three ESKAPE pathogens: *E. faecalis*, *P. aeruginosa*, and *E. cloacae*.

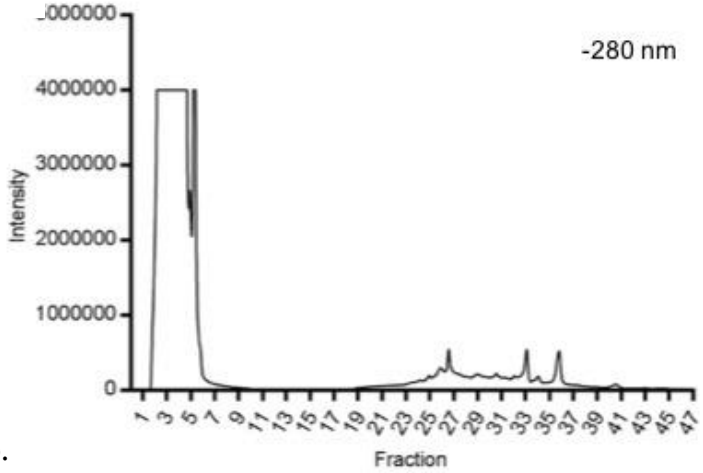
6.6 Conclusion

Using PepSAVI-MS, we begin to explore the antimicrobial activity from the aerial tissue of *A. tricolor*. Bioactivity results indicate strong activity against the gram-negative bacterium *Escherichia coli*, which is retained through both SCX and reversed phase fractionation. Using the PepSAVIms statistical analysis package, the top 20 candidates most-likely contributing to the observed bioactivity have been identified, and further work is being performed to identify the specific bioactive component. The purified bioactive peptide will undergo activity validation against *E. coli* and will be screened for activity against a full panel of gram-negative and gram-positive bacteria to assess full bioactivity potential. This work represents the first novel peptide identified via PepSAVI-MS and showcases the potential of this pipeline for natural product bioactive peptide discovery.

6.7 Figures

Figure 6.1 *Amaranthus tricolor* plant growth at 2, 4, and 8 weeks after planting the seeds. The picture at 8 weeks was taken directly before harvesting.

a.



b.

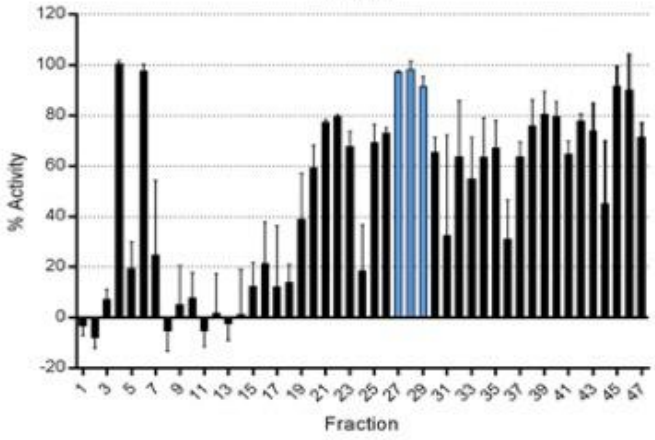


Figure 6.2 *A. tricolor* SCX peptide library (a) 280 nm UV trace and (b) bioactivity profile vs. *E. coli*.

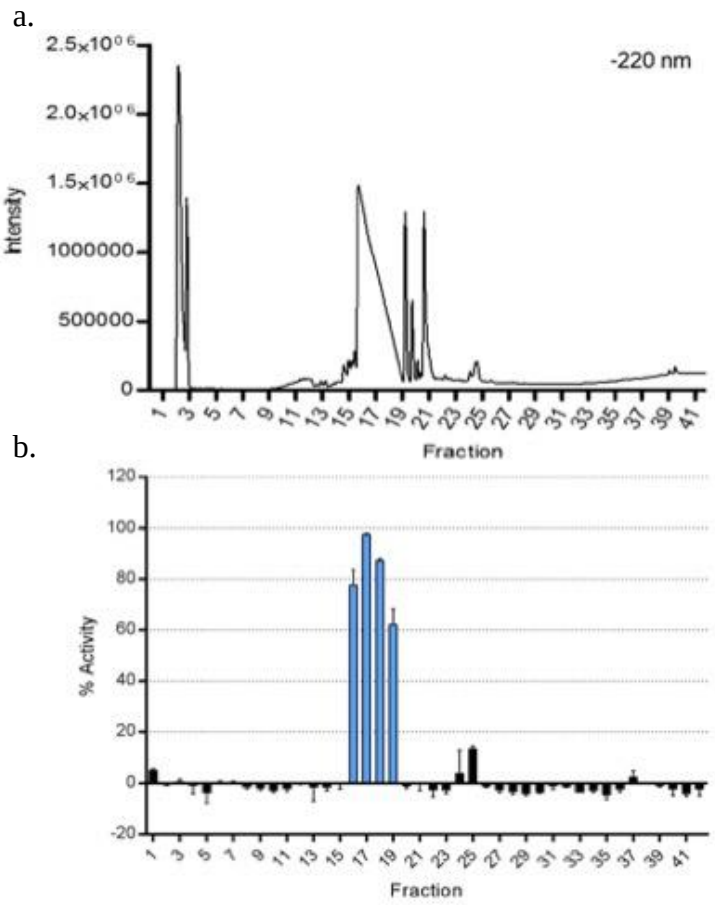


Figure 6.3 *A. tricolor* RP peptide library (a) 220 nm UV trace and (b) bioactivity profile vs. *E. coli*.

Rank	*m/z*	Charge	Mass
1	568.948	3	1703.821
2	385.365	5	1921.783
3	405.574	3	1213.699
4	360.306	9	3233.679
5	404.823	5	2019.078
6	641.588	4	2562.319
7	725.958	3	2174.851
8	619.566	4	2474.231
9	510.281	3	1527.819
10	405.308	4	1617.202
11	385.895	3	1154.663
12	495.850	5	2474.211
13	628.864	2	1255.712
14	404.323	4	1613.259
15	874.924	2	1747.833
16	558.645	3	1672.912
17	394.973	4	1575.861
18	425.249	4	1696.965
19	582.726	5	2908.592
20	543.814	4	2171.227

Figure 6.4 *A. tricolor* top 20 compounds after statistical modeling of RP peptide library.

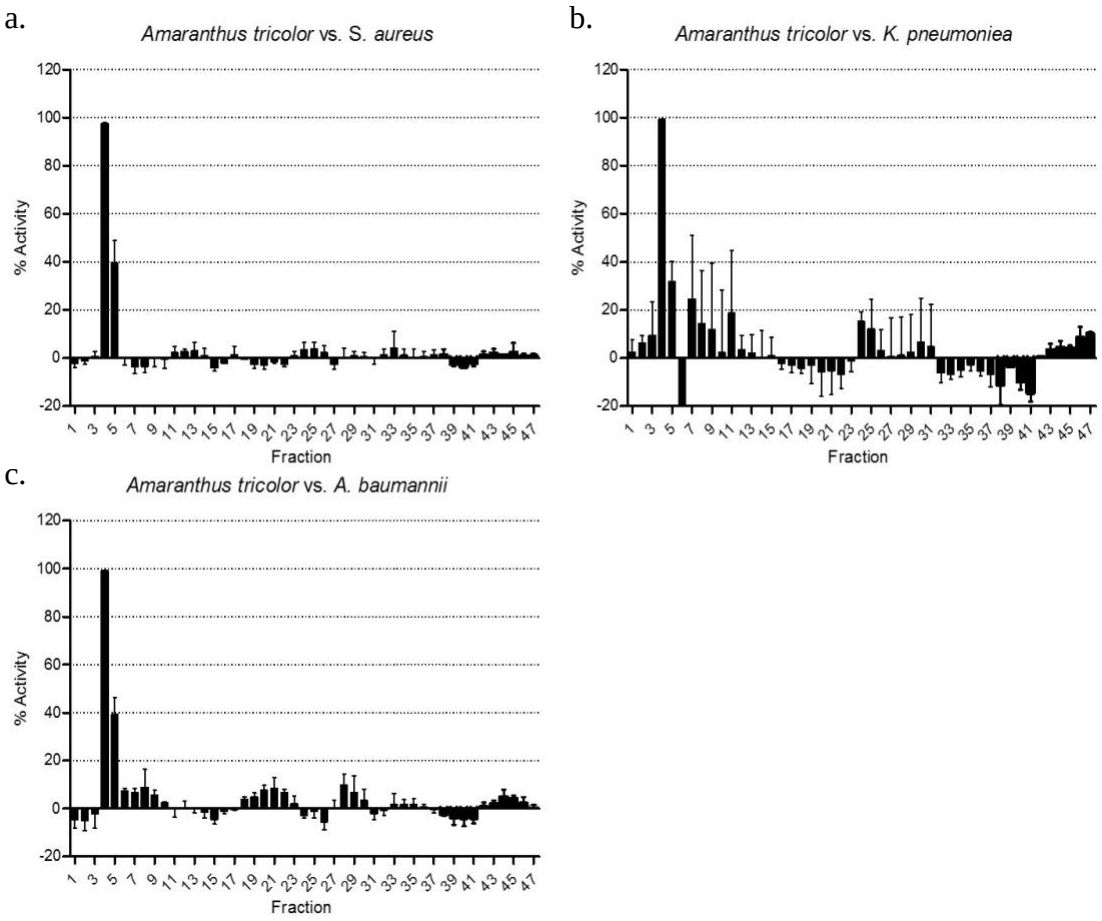


Figure 6.5 *A. tricolor* SCX peptide library bioactivity against a subset of the ESKAPE pathogens including (a) *Staphylococcus aureus*, (b) *Klebsiella pneumoniae*, and (c) *Acinetobacter baumannii*.

REFERENCES

1. Peter, K.; Gandhi, P., Rediscovering the therapeutic potential of Amaranthus species: A review. *Egyptian Journal of Basic and Applied Sciences* **2017,** *4* (3), 196-205.

2. Baral, M.; Datta, A.; Chakraborty, S.; Chakraborty, P., Pharmacognostic studies on stem and leaves of Amaranthus spinosus Linn. *International Journal of Applied Biology and Pharmaceutical Technology* **2011,** *2* (1), 41-47.

3. Agra, M. D. F.; Silva, K. N.; Basílio, I. J. L. D.; De Freitas, P. F.; Barbosa-Filho, J. M., Survey of medicinal plants used in the region Northeast of Brazil. *Brazilian Journal of Pharmacognosy* **2008,** *18* (3), 472-508.

4. Maldonado-Cervantes, E.; Jeong, H. J.; León-Galván, F.; Barrera-Pacheco, A.; De León-Rodríguez, A.; González De Mejia, E.; De Lumen, B. O.; Barba De La Rosa, A. P., Amaranth lunasin-like peptide internalizes into the cell nucleus and inhibits chemical carcinogen-induced transformation of NIH-3T3 cells. *Peptides* **2010,** *31* (9), 1635-1642.

5. Silva-Sanchez, C.; de la Rosa, A. P.; Leon-Galvan, M. F.; de Lumen, B. O.; de Leon-Rodriguez, A.; de Mejia, E. G., Bioactive peptides in amaranth (Amaranthus hypochondriacus) seed. *J. Agric. Food. Chem.* **2008,** *56* (4), 1233-1240.

6. Tiengo, A.; Faria, M.; Netto, F. M., Characterization and ACE-inhibitory activity of amaranth proteins. *J. Food Sci.* **2009,** *74* (5), H121-126.

7. de la Rosa, A. P.; Montoya, A. B.; Martinez-Cuevas, P.; Hernandez-Ledesma, B.; Leon-Galvan, M. F.; De Leon-Rodriguez, A.; Gonzalez, C., Tryptic amaranth glutelin digests induce endothelial nitric oxide production through inhibition of ACE: antihypertensive role of amaranth peptides. *Nitric Oxide* **2010,** *23* (2), 106-111.

8. Barrio, D. A.; Anon, M. C., Potential antitumor properties of a protein isolate obtained from the seeds of Amaranthus mantegazzianus. *Eur. J. Nutr.* **2010,** *49* (2), 73-82.

9. Chopra, R. N.; Nayar, S. L.; Chopra, I. C.; Asolkar, L. V.; Kakkar, K. K.; Chakre, O. J.; Varma, B. S.; Council of, S.; Industrial, R., *Glossary of Indian medicinal plants ; [with] Supplement*. Council of Scientific & Industrial Research: New Delhi, **1956**.

10. Biswas, B.; Rogers, K.; McLaughlin, F., et al. Antimicrobial activities of leaf extracts of Guava (Psidium guajava L.) on two gram-negative and gram-positive bacteria. Int J Microbiol 2013. *Article ID* **2013**, 746165.

11. Bihani, G. V.; Bodhankar, S. L.; Kadam, P. P.; Zambare, G. N., Anti-nociceptive and anti-inflammatory activity of hydroalcoholic extract of leaves of amaranthus tricolor L. *Der Pharmacia Lettre* **2013,** *5* (3), 48-55.

12. Devaraj, V. C.; Krishna, B., Antiulcer activity of a polyherbal formulation (PHF) from Indian medicinal plants. *Chinese Journal of Natural Medicines* **2013,** *11* (2), 145-148.

13. Devaraj, V. C.; Krishna, B. G., Gastric antisecretory and cytoprotective effects of leaf extracts of Amaranthus tricolor Linn, in rats. *Journal of Chinese Integrative Medicine* **2011,** *9* (9), 1031-1038.

14. Al-Dosari, M. S., The effectiveness of ethanolic extract of Amaranthus tricolor L.: A natural hepatoprotective agent. *Am. J. Chin. Med.* **2010,** *38* (6), 1051-1064.

15. Rahmatullah, M.; Hosain, M.; Rahman, S.; Rahman, S.; Akter, M.; Rahman, F.; Rehana, F.; Munmun, M.; Kalpana, M. A., Antihyperglycaemic and antinociceptive activity evaluation of methanolic extract of whole plant of Amaranthus tricolour L. (Amaranthaceae). *African journal of traditional, complementary, and alternative medicines : AJTCAM / African Networks on Ethnomedicines* **2013,** *10* (5), 408-411.

16. Amornrit, W.; Santiyanont, R., Neuroprotective effect of Amaranthus lividus and Amaranthus tricolor and their effects on gene expression of RAGE during oxidative stress in SH-SY5Y cells. *Genet. Mol. Res.* **2016,** *15* (2), 1-14.

17. Kushwaha, S.; Chawla, P.; Kochhar, A., Effect of supplementation of drumstick (Moringa oleifera) and amaranth (Amaranthus tricolor) leaves powder on antioxidant profile and oxidative status among postmenopausal women. *J. Food Sci. Technol.* **2014,** *51* (11), 3464-3469.

18. Kirkpatrick, C. L.; Broberg, C. A.; McCool, E. N.; Lee, W. J.; Chao, A.; McConnell, E. W.; Pritchard, D. A.; Hebert, M.; Fleeman, R.; Adams, J.; Jamil, A.; Madera, L.; Stromstedt, A. A.; Goransson, U.; Liu, Y.; Hoskin, D. W.; Shaw, L. N.; Hicks, L. M., The "PepSAVI-MS" Pipeline for Natural Product Bioactive Peptide Discovery. *Anal. Chem.* **2017,** *89* (2), 1194-1201.

19. Fleeman, R.; LaVoi, T. M.; Santos, R. G.; Morales, A.; Nefzi, A.; Welmaker, G. S.; Medina-Franco, J. L.; Giulianotti, M. A.; Houghten, R. A.; Shaw, L. N., Combinatorial Libraries As a Tool for the Discovery of Novel, Broad-Spectrum Antibacterial Agents Targeting the ESKAPE Pathogens. *J. Med. Chem.* **2015,** *58* (8), 3340-3355.

20. Johnson, M. T.; Carpenter, E. J.; Tian, Z.; Bruskiewich, R.; Burris, J. N.; Carrigan, C. T.; Chase, M. W.; Clarke, N. D.; Covshoff, S.; Depamphilis, C. W.; Edger, P. P.; Goh, F.; Graham, S.; Greiner, S.; Hibberd, J. M.; Jordon-Thaden, I.; Kutchan, T. M.; Leebens-Mack, J.; Melkonian, M.; Miles, N.; Myburg, H.; Patterson, J.; Pires, J. C.; Ralph, P.; Rolf, M.; Sage, R. F.; Soltis, D.; Soltis, P.; Stevenson, D.; Stewart, C. N., Jr.; Surek, B.; Thomsen, C. J.; Villarreal, J. C.; Wu, X.; Zhang, Y.; Deyholos, M. K.; Wong, G. K., Evaluating methods for isolating total RNA and predicting the success of sequencing phylogenetically diverse plant transcriptomes. *PLoS One* **2012,** *7* (11), e50226.

21. Matasci, N.; Hung, L. H.; Yan, Z.; Carpenter, E. J.; Wickett, N. J.; Mirarab, S.; Nguyen, N.; Warnow, T.; Ayyampalayam, S.; Barker, M.; Burleigh, J. G.; Gitzendanner, M. A.; Wafula, E.; Der, J. P.; dePamphilis, C. W.; Roure, B.; Philippe, H.; Ruhfel, B. R.; Miles, N. W.; Graham, S. W.; Mathews, S.; Surek, B.; Melkonian, M.; Soltis, D. E.; Soltis, P. S.; Rothfels, C.; Pokorny, L.; Shaw, J. A.; DeGironimo, L.; Stevenson, D. W.; Villarreal,

J. C.; Chen, T.; Kutchan, T. M.; Rolf, M.; Baucom, R. S.; Deyholos, M. K.; Samudrala, R.; Tian, Z.; Wu, X.; Sun, X.; Zhang, Y.; Wang, J.; Leebens-Mack, J.; Wong, G. K., Data access for the 1,000 Plants (1KP) project. *Gigascience* **2014,** *3* (17), 1-10.

22. Wickett, N. J.; Mirarab, S.; Nguyen, N.; Warnow, T.; Carpenter, E.; Matasci, N.; Ayyampalayam, S.; Barker, M. S.; Burleigh, J. G.; Gitzendanner, M. A.; Ruhfel, B. R.; Wafula, E.; Der, J. P.; Graham, S. W.; Mathews, S.; Melkonian, M.; Soltis, D. E.; Soltis, P. S.; Miles, N. W.; Rothfels, C. J.; Pokorny, L.; Shaw, A. J.; DeGironimo, L.; Stevenson, D. W.; Surek, B.; Villarreal, J. C.; Roure, B.; Philippe, H.; dePamphilis, C. W.; Chen, T.; Deyholos, M. K.; Baucom, R. S.; Kutchan, T. M.; Augustin, M. M.; Wang, J.; Zhang, Y.; Tian, Z.; Yan, Z.; Wu, X.; Sun, X.; Wong, G. K.; Leebens-Mack, J., Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc. Natl. Acad. Sci. U. S. A.* **2014,** *111* (45), E4859-4868.

23. Xie, Y.; Wu, G.; Tang, J.; Luo, R.; Patterson, J.; Liu, S.; Huang, W.; He, G.; Gu, S.; Li, S.; Zhou, X.; Lam, T. W.; Li, Y.; Xu, X.; Wong, G. K.; Wang, J., SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. *Bioinformatics* **2014,** *30* (12), 1660-1666.

CHAPTER 7: Future Work and Conclusions

7.1 Future Work

7.1.1 High-Throughput Screening of Traditional Medicinal Plants

High-throughput screening of traditional medicinal plants is continuously underway in the Hicks Laboratory. Plant species have been selected based on criteria including known biological activity, traditional use, and availability. To date, almost 70 botanical species have been grown and harvested with 51 crude extracts and 31 peptide libraries created for high-throughput screening of medicinal plant species. Peptide libraries will be screened against a panel of pathogen selected based on literature precedent. The most promising plant species will be identified and prioritized based on strength and breadth of activity. Prioritized species will be subject to LC-MS/MS and statistical analysis for determination of putative bioactive peptide targets. Selected peptides from the top 20 compound lists will undergo subsequent characterization and activity validation for novel bioactive peptide discovery.

7.1.2 Induced Expression of Bioactive Peptides

Although many bioactive peptides are constitutively expressed, expression of bioactive peptides can also be initiated or increased through biotic and abiotic stress conditions. Incorporation of induced bioactive peptides into the initial peptide library allows for access to peptides that might otherwise be unavailable. Abiotic stress conditions, including salt, wounding, and chemical stress, will be the initial focus of induced bioactive peptides due to ease of implementation. To introduce abiotic stress conditions, hydroponic and liquid culture growth methods have been developed. Additional methods involving wounding have been established.

Regardless of induction strategy, after the initial stress is applied plant material is harvested and taken through the PepSAVI-MS pipeline without any further modification.

7.1.3 Validation Challenges

The developed PepSAVI-MS pipeline rapidly identifies bioactive peptide targets that require sequence identification, structural characterization, and activity validation. Although prioritization of these compounds is a necessary first step, significant focus on peptide targets will be required to fully characterize a bioactive peptide. While these steps are beyond the scope of this dissertation, they are crucial to complete the bioactive peptide discovery process. As such, the remaining sections identify possible challenges and ideas to address them.

7.1.4 Peptide Isolation/Synthesis

Target peptides must be isolated or synthesized for activity validation and full structural characterization. The overall complexity of the peptide, accounting for primary sequence, secondary structure, and complexity of post-translational modifications, as well as the abundance of the target peptide produced by the natural product, will dictate the appropriate route forward for isolation. Isolation options include purification from the native plant species if sufficient material can be obtained[1] and solid phase chemical synthesis[2-4]. To enable production of peptides and small proteins, Fmoc-based solid-phase peptide synthesis (SPPS) will be performed using a flow chemistry-based platform[5-6]. Using this proven instrument, peptide fragments are rapidly generated and further purified with modern chromatography methods. In order to access linear structures greater than 50 amino acids, peptide fragments are joined together using established *in situ* native-chemical ligation methodology, where the activated *C*-terminus of one peptide is chemoselectively joined with another peptide fragment containing an *N*-terminal Cys residue[7]. Strategies to form backbone-cyclized sequences will utilize similar literature precedent under

high dilution conditions[8] or enzyme-mediated processes for backbone cyclization[9]. Towards the synthesis of disulfide-containing peptide targets, the wealth of prior experimentation will serve as guide, where careful consideration of buffer additives can facilitate improved folding[10]. Initial conditions to afford correct disulfide regiochemistry will include dilute iodine or DMSO as oxidants; however, use of thermodynamic folding conditions with thiol-based redox pairs such as GSH/GSSG may better mimic endogenous conditions[8, 11]. In cases where these oxidative folding methods fail, orthogonal *S*-Cys protection will be employed to engineer desired disulfide pairing[12].

7.1.5 MOA Determination

Mechanisms of action (MOAs) of high priority peptides will be explored using a two-pronged strategy. First, compounds will be assessed for common antibacterial MOAs, using the bacterial cytological profiling (BCP) method pioneered by the Pogliano group at the University of California San Diego[13]. BCP uses fluorescent microscopy to assess alterations in bacterial ultrastructure upon exposure to antibacterial agents. By comparison to the discrete and specific effects of known classes of antibacterial agents, one can begin to develop an understanding of novel drug MOA. This approach has been successful in identifying the MOA of a number of novel antibacterial agents[13-19], including the identification of a novel class of bis-cyclic guanidines that affect bacteria via the inhibition of protein synthesis. Novel leads that meet progression criteria will first be assessed using BCP methodologies.

A second, complimentary approach, employing CETSA (Cellular Thermal Shift Assays) will be used for compounds requiring further analysis[20]. CETSA works on the basis that ligands (drugs) bound to target proteins increase their stability; thus, when one applies heat to treated and untreated samples, followed by mass spectrometric analysis, a shift in thermal stability is

observed for proteins bound by the compound in question. This approach has been used extensively in the context of eukaryotic drug discovery, particular in the area of oncology[21-24].

7.1.6 Translation

Upon successful characterization of active peptide targets, strategies to improve *in vivo* stability and immunogenicity profiles will be investigated to assess translation potential. Such strategies include incorporation of D-amino acids, amino acid substitution with non-native peptide sidechains, *N*-methylation of peptide backbone, incorporation of peptoid or beta-amino acids, disulfide bond replacement, and peptide stapling[25-28]. As a final measure of effectivity for lead peptides, *in vivo* efficacy testing will be performed using lethal murine peritonitis models of infection, which have been extensively described in the context of drug discovery for each of the ESKAPE organisms[29-34].

7.2 Conclusions

PepSAVI-MS is a powerful and versatile pipeline for natural product bioactive peptide discovery. Successful identification of cyO2 (*Viola odorata*), KP4 (*Ustilago maydis*), and Bac-21 (*Enterococcus faecalis*) demonstrate proof-of-principle for the developed pipeline and establish its utility across source material from plants, fungi, and bacteria. PepSAVI-MS is easily adaptable to accommodate a variety of bioassay screens, including antibacterial, antifungal, and anticancer, in both liquid- and agar-based formats. After bioactivity screening, peptide libraries can be analyzed via LC-MS intact or digested (for genome sequenced organisms), further extending the utility of the pipeline to use with resolving power-limited mass spectrometers. The established pipeline is being applied for novel bioactive peptide discovery and has resulted in many promising leads that are currently being pursued further.

In comparison to current methods for targeted bioactive peptide discovery, PepSAVI-MS provides a direct measure of bioactivity while remaining high-throughput and requires no *a priori* knowledge of an organism's genome or the biosynthetic pathways of known bioactive peptides. Combination of PepSAVI-MS with other peptidomics tools, including genome mining and molecular networking, have the possibility to greatly expand its capabilities when this information is available. PepSAVI-MS is uniquely poised probe natural product extracts with a new lens and has the potential to expedite the search for novel bioactive peptides.

REFERENCES

1. Pranting, M.; Loov, C.; Burman, R.; Goransson, U.; Andersson, D. I., The cyclotide cycloviolacin O2 from Viola odorata has potent bactericidal activity against Gram-negative bacteria. *J Antimicrob Chemoth* **2010,** *65* (9), 1964-1971.

2. Leta Aboye, T.; Clark, R. J.; Craik, D. J.; Goransson, U., Ultra-stable peptide scaffolds for protein engineering-synthesis and folding of the circular cystine knotted cyclotide cycloviolacin O2. *Chembiochem : a European journal of chemical biology* **2008,** *9* (1), 103-113.

3. Thongyoo, P.; Tate, E. W.; Leatherbarrow, R. J., Total synthesis of the macrocyclic cysteine knot microprotein MCoTI-II. *Chemical communications* **2006,** (27), 2848-2850.

4. Daly, N. L.; Love, S.; Alewood, P. F.; Craik, D. J., Chemical synthesis and folding pathways of large cyclic polypeptides: studies of the cystine knot polypeptide kalata B1. *Biochemistry* **1999,** *38* (32), 10606-10614.

5. Simon, M. D.; Heider, P. L.; Adamo, A.; Vinogradov, A. A.; Mong, S. K.; Li, X.; Berger, T.; Policarpo, R. L.; Zhang, C.; Zou, Y.; Liao, X.; Spokoyny, A. M.; Jensen, K. F.; Pentelute, B. L., Rapid flow-based peptide synthesis. *Chembiochem : a European journal of chemical biology* **2014,** *15* (5), 713-720.

6. Mijalis, A. J.; Thomas, D. A., 3rd; Simon, M. D.; Adamo, A.; Beaumont, R.; Jensen, K. F.; Pentelute, B. L., A fully automated flow-based approach for accelerated peptide synthesis. *Nature chemical biology* **2017,** *13* (5), 464-466.

7. Zheng, J. S.; Tang, S.; Qi, Y. K.; Wang, Z. P.; Liu, L., Chemical synthesis of proteins using peptide hydrazides as thioester surrogates. *Nature protocols* **2013,** *8* (12), 2483-2495.

8. Tam, J. P.; Lu, Y. A.; Yu, Q. T., Thia zip reaction for synthesis of large cyclic peptides: Mechanisms and applications. *Journal of the American Chemical Society* **1999,** *121* (18), 4316-4324.

9. Harris, K. S.; Durek, T.; Kaas, Q.; Poth, A. G.; Gilding, E. K.; Conlan, B. F.; Saska, I.; Daly, N. L.; van der Weerden, N. L.; Craik, D. J.; Anderson, M. A., Efficient backbone cyclization of linear peptides by a recombinant asparaginyl endopeptidase. *Nat Commun* **2015,** *6,* 10199.

10. Wong, C. T. T.; Taichi, M.; Nishio, H.; Nishiuchi, Y.; Tam, J. P., Optimal Oxidative Folding of the Novel Antimicrobial Cyclotide from Hedyotis biflora Requires High Alcohol Concentrations. *Biochemistry* **2011,** *50* (33), 7275-7283.

11. Wu, Z.; Ericksen, B.; Tucker, K.; Lubkowski, J.; Lu, W., Synthesis and characterization of human alpha-defensins 4-6. *The journal of peptide research : official journal of the American Peptide Society* **2004,** *64* (3), 118-125.

12. Yang, Y.; Sweeney, W. V.; Schneider, K.; Chait, B. T.; Tam, J. P., Two-step selective formation of three disulfide bridges in the synthesis of the C-terminal epidermal growth factor-like domain in human blood coagulation factor IX. *Protein science : a publication of the Protein Society* **1994,** *3* (8), 1267-1275.

13. Nonejuie, P.; Burkart, M.; Pogliano, K.; Pogliano, J., Bacterial cytological profiling rapidly identifies the cellular pathways targeted by antibacterial molecules. *Proceedings of the National Academy of Sciences of the United States of America* **2013,** *110* (40), 16169-16174.

14. Araujo-Bazan, L.; Ruiz-Avila, L. B.; Andreu, D.; Huecas, S.; Andreu, J. M., Cytological Profile of Antibacterial FtsZ Inhibitors and Synthetic Peptide MciZ. *Frontiers in microbiology* **2016,** *7* (5558), 1-17.

15. Lamsa, A.; Liu, W. T.; Dorrestein, P. C.; Pogliano, K., The Bacillus subtilis cannibalism toxin SDP collapses the proton motive force and induces autolysis. *Molecular microbiology* **2012,** *84* (3), 486-500.

16. Lamsa, A.; Lopez-Garrido, J.; Quach, D.; Riley, E. P.; Pogliano, J.; Pogliano, K., Rapid Inhibition Profiling in Bacillus subtilis to Identify the Mechanism of Action of New Antimicrobials. *ACS chemical biology* **2016,** *11* (8), 2222-2231.

17. Nayar, A. S.; Dougherty, T. J.; Ferguson, K. E.; Granger, B. A.; McWilliams, L.; Stacey, C.; Leach, L. J.; Narita, S.; Tokuda, H.; Miller, A. A.; Brown, D. G.; McLeod, S. M., Novel antibacterial targets and compounds revealed by a high-throughput cell wall reporter assay. *Journal of bacteriology* **2015,** *197* (10), 1726-1734.

18. Nonejuie, P.; Trial, R. M.; Newton, G. L.; Lamsa, A.; Ranmali Perera, V.; Aguilar, J.; Liu, W. T.; Dorrestein, P. C.; Pogliano, J.; Pogliano, K., Application of bacterial cytological profiling to crude natural product extracts reveals the antibacterial arsenal of Bacillus subtilis. *The Journal of antibiotics* **2016,** *69* (5), 353-361.

19. Pogliano, J.; Pogliano, N.; Silverman, J. A., Daptomycin-mediated reorganization of membrane architecture causes mislocalization of essential cell division proteins. *Journal of bacteriology* **2012,** *194* (17), 4494-4504.

20. Molina, D. M.; Jafari, R.; Ignatushchenko, M.; Seki, T.; Larsson, E. A.; Dan, C.; Sreekumar, L.; Cao, Y. H.; Nordlund, P., Monitoring Drug Target Engagement in Cells and Tissues Using the Cellular Thermal Shift Assay. *Science* **2013,** *341* (6141), 84-87.

21. Alshareef, A.; Zhang, H. F.; Huang, Y. H.; Wu, C. S.; Zhang, J. D.; Wang, P.; El-Sehemy, A.; Fares, M.; Lai, R., The use of cellular thermal shift assay (CETSA) to study Crizotinib resistance in ALK-expressing human cancers. *Scientific reports* **2016,** *6*, 33710.

22. Huang, R.; Ayine-Tora, D. M.; Muhammad Rosdi, M. N.; Li, Y.; Reynisson, J.; Leung, I. K., Virtual screening and biophysical studies lead to HSP90 inhibitors. *Bioorg Med Chem Lett* **2017,** *27* (2), 277-281.

23. Qin, D.; Wang, W. W.; Lei, H.; Luo, H.; Cai, H. Y.; Tang, C. X.; Wu, Y. Z.; Wang, Y. Y.; Jin, J.; Xiao, W. L.; Wang, T. D.; Ma, C. M.; Xu, H. Z.; Zhang, J. F.; Gao, F. H.; Wu, Y. L., CDDO-Me reveals USP7 as a novel target in ovarian cancer cells. *Oncotarget* **2016,** *7* (47), 77096-77109.

24. Warpman Berglund, U.; Sanjiv, K.; Gad, H.; Kalderen, C.; Koolmeister, T.; Pham, T.; Gokturk, C.; Jafari, R.; Maddalo, G.; Seashore-Ludlow, B.; Chernobrovkin, A.; Manoilov, A.; Pateras, I. S.; Rasti, A.; Jemth, A. S.; Almlof, I.; Loseva, O.; Visnes, T.; Einarsdottir, B. O.; Gaugaz, F. Z.; Saleh, A.; Platzack, B.; Wallner, O. A.; Vallin, K. S.; Henriksson, M.; Wakchaure, P.; Borhade, S.; Herr, P.; Kallberg, Y.; Baranczewski, P.; Homan, E. J.; Wiita, E.; Nagpal, V.; Meijer, T.; Schipper, N.; Rudd, S. G.; Brautigam, L.; Lindqvist, A.; Filppula, A.; Lee, T. C.; Artursson, P.; Nilsson, J. A.; Gorgoulis, V. G.; Lehtio, J.; Zubarev, R. A.; Scobie, M.; Helleday, T., Validation and development of MTH1 inhibitors for treatment of cancer. *Annals of oncology : official journal of the European Society for Medical Oncology* **2016,** *27* (12), 2275-2283.

25. Chatterjee, J.; Gilon, C.; Hoffman, A.; Kessler, H., N-methylation of peptides: a new perspective in medicinal chemistry. *Accounts of chemical research* **2008,** *41* (10), 1331-1342.

26. Lau, Y. H.; de Andrade, P.; Wu, Y.; Spring, D. R., Peptide stapling techniques based on different macrocyclisation chemistries. *Chemical Society reviews* **2015,** *44* (1), 91-102.

27. Binder, J. B.; Raines, R. T., Olefin metathesis for chemical biology. *Current opinion in chemical biology* **2008,** *12* (6), 767-773.

28. Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H., Efficient Method for the Preparation of Peptoids [Oligo(N-Substituted Glycines)] by Submonomer Solid-Phase Synthesis. *Journal of the American Chemical Society* **1992,** *114* (26), 10646-10647.

29. Mimoz, O.; Gregoire, N.; Poirel, L.; Marliat, M.; Couet, W.; Nordmann, P., Broad-spectrum beta-lactam antibiotics for treating experimental peritonitis in mice due to Klebsiella pneumoniae producing the carbapenemase OXA-48. *Antimicrobial agents and chemotherapy* **2012,** *56* (5), 2759-2760.

30. Li, G. Q.; Bai, X. G.; Li, C. R.; Yang, X. Y.; Hu, X. X.; Yuan, M.; Zhang, W. X.; Lou, R. H.; Guo, H. Y.; Jiang, J. D.; You, X. F., In vivo antibacterial activity of chinfloxacin, a new fluoroquinolone antibiotic. *The Journal of antimicrobial chemotherapy* **2012,** *67* (4), 955-961.

31. Pucci, M. J.; Podos, S. D.; Thanassi, J. A.; Leggio, M. J.; Bradbury, B. J.; Deshpande, M., In vitro and in vivo profiles of ACH-702, an isothiazoloquinolone, against bacterial pathogens. *Antimicrobial agents and chemotherapy* **2011,** *55* (6), 2860-2871.

32. Im, W. B.; Choi, S. H.; Park, J. Y.; Choi, S. H.; Finn, J.; Yoon, S. H., Discovery of torezolid as a novel 5-hydroxymethyl-oxazolidinone antibacterial agent. *European journal of medicinal chemistry* **2011,** *46* (4), 1027-1039.

33. Page, M. G.; Dantier, C.; Desarbre, E.; Gaucher, B.; Gebhardt, K.; Schmitt-Hoffmann, A., In vitro and in vivo properties of BAL30376, a beta-lactam and dual beta-lactamase inhibitor combination with enhanced activity against Gram-negative Bacilli that express multiple beta-lactamases. *Antimicrobial agents and chemotherapy* **2011,** *55* (4), 1510-1519.

34. Song, J. H.; Ko, K. S.; Suh, J. Y.; Oh, W. S.; Kang, C. I.; Chung, D. R.; Peck, K. R.; Lee, N. Y.; Lee, W. G., Clinical implications of vancomycin-resistant Enterococcus faecium (VRE) with VanD phenotype and vanA genotype. *The Journal of antimicrobial chemotherapy* **2008,** *61* (4), 838-844.